二十世纪西方哲学译丛

实验哲学导论

EXPERIMENTAL PHILOSOPHY: AN INTRODUCTION

[美] 约书亚·亚历山大 著
楼 巍 译

上海译文出版社

VISIT…

LanzaroteCaliente.com

世纪出版

上架建议：外国哲学

ISBN 978-7-5327-6344-3

9 787532 763443 >

定 价： 35.00 元

易文网：www.ewen.cc

二十世纪西方哲学译丛

实验哲学导论

EXPERIMENTAL PHILOSOPHY: AN INTRODUCTION

[美] 约书亚·亚历山大 著
楼 巍 译



上海译文出版社

图书在版编目(CIP)数据

实验哲学导论/(美)亚历山大(Alexander,J.)著;楼巍译.
—上海:上海译文出版社,2013.11
(二十世纪西方哲学译丛)
书名原文:Experimental Philosophy:An Introduction
ISBN 978-7-5327-6344-3

Ⅰ.实… Ⅱ.①亚… ②楼… Ⅲ①哲学 Ⅳ.①B

中国版本图书馆 CIP 数据核字(2013)第 171539 号

Experimental Philosophy: An Introduction
by Joshua Alexander
Copyright © Polity Press 2012
This Edition Is Published By Arrangement With Polity Press Ltd., Cambridge
All Rights Reserved.
No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

图字:09-2005-551 号

本书中文简体字专有出版权
归本社独家所有,非经本社同意不得转载、摘编或复制

实验哲学导论
〔美〕约书亚·亚历山大 著
楼 巍 译

上海世纪出版股份有限公司
上海译文出版社出版
网址:www.yiwen.com.cn
上海世纪出版股份有限公司发行中心发行
200001 上海福建中路 193 号 www.ewen.cc
上海商务联西印刷有限公司印刷

开本 890×1240 1/32 印张 6.5 插页 4 字数 165,000
2013 年 11 月第 1 版 2013 年 11 月第 1 次印刷
ISBN 978-7-5327-6344-3/B·370
定价:35.00 元

如有质量问题,请与承印厂质量科联系。T:021-56135113

实验哲学导论

EXPERIMENTAL PHILOSOPHY: AN INTRODUCTION

本书得到以下课题资助

教育部人文社科青年基金项目“最后阶段的维特根斯坦与知识论”（11YJC720029）
中央高校基本科研业务费专项资金资助项目“实验哲学”（2012221003）
教育部新世纪优秀人才支持计划“实验知识论”（NCET-11-0290）

《二十世纪西方哲学译丛》选收二十世纪西方哲学界各主要流派影响较大的著作。通过有选择的译介，旨在增进文化积累，拓展学术视野，丰富研究课题，为了解和研讨现代西方哲学提供系统而完整的第一手资料，以利于我国理论界、学术界深化对西方文化的借鉴和批判。

中文版序言

本书考察的是哲学的认知，考察我们的心灵是如何工作的，以及我们是如何思考哲学问题的，其目的在于更好地理解应该由什么样的规范来支配我们的认知行为以及我们应该采用什么样的认知方法。它勾勒了一个正在壮大中的运动，这场运动的局面在不断地变化。在中文版序言中，我想简要刻画一下实验哲学三个最近的进展。这三个进展仅仅呈现了自本书去年夏天首次出版以来实验哲学领域的一些新的重要工作的一小部分。

知识和信念

这里有一个关于知识的普遍信念：一个人不可能知道某些她不相信的东西。但是考虑一下如下这个小场景：

> 凯特正在参加历史考试。她学习很认真，而且到目前为止所有题目她都答得挺好。现在她遇到了最后一道题，是这样的："伊丽莎白女王是哪一年去世的？"看到这个问题，凯特觉得挺轻松，因为她已经预测到这道题并且记住了答案。凯特正要停下来回想一下时间，老师打了个岔，宣布只剩下一分钟了。现在凯特慌了。她的手牢牢地握着笔。她的大脑一片空白，她什么也想不起来。她觉得只能瞎猜了。因此，她沮丧而心神不宁地写下了"1603 年"——当

然，这恰好就是正确答案。

令人惊奇的是，布莱克·迈尔斯-舒尔茨（Blake Myers-Schulz）和埃里克·史怀泽格贝尔（Eric Schwitzgebel）（即将出版）发现，比起认为凯特相信伊丽莎白女王死于1603年，普通人更愿意认为凯特知道伊丽莎白女王死于1603年。搞清楚这些令人惊奇的结论有助于揭示一些针对我们的信念概念的新洞见。比如，戴维·罗斯（David Rose）和乔纳森·谢弗（Jonathan Schaffer）（即将出版）用这些结论来强调在当下信念（occurrent belief）和性状信念（dispositional belief）之间的重要区分，他们提出凯特无法拥有相关的当下信念，因为她的性状信念被老师的通告所带来的压力掩盖住了。韦斯利·布克瓦尔特（Wesley Buckwalter）、戴维·罗斯和约翰·图利（John Turri）（即将出版）提出，这些结论让我们有理由相信存在着关于信念的厚观念和薄观念，薄的信念仅仅是命题性态度（这些态度将某些事态反映为是真的），而厚的信念则是我们积极地支持或促进的那种认知状态。这些是一些很不错的进展，特别是因为它们对于当代知识论中的一些其他的论题（包括关于信念规范的争论、信念的伦理学、知识和断定之间的关系）有着重要的意义。

知识和断定

某人说了一些什么，而又补充说她不知道这是真的，这似乎是错误的（摩尔，1962年）。大多数哲学家接受这条规则或类似这条规则的东西，我们应该只断定我们知道的东西。但是，针对这条规则，似乎存在着一些反例，在这些例子中，某个行动者恰当地断定了某个东西而不相信它是真的，或者这些例子涉及这

样的情况，在这里尽管有着充分证据，但最后发现这个东西是假的（莱奇[Lackey]，2005年）。这些反例将会为任何一种对断定的知识性考察带来严重的问题，因此值得注意的是在这些假定的反例中至少有一些反例最终是反直觉的（图利，即将出版）。结果是，普通人认为我们不应该去断定错误的主张，即使那些主张得到了所有可资利用的证据的充分支持，而且可以用我们如下这一倾向来将至少一些假定的反例解释掉（explain away）：我们会倾向于一开始就错误地否定一个违法行为的发生，以此来宽恕那种不应该受到指责的违法行为。这为如下看法提供了进一步的证据，即断定的规范毕竟还是事实性的（factive）。同时，对于我们日常的语言实践（包括知识归属的实践）而言，这有着重要的意义。

知识归属、显著性和知识的咒语

在当代知识论中存在着一个持久的争论，那就是知识归属的标准是否取决于在一个给定的会话语境（conversational context）中什么样的可能性被搞成是显著的。使得这场争论变得尤其有趣的是争论双方都同意，比起那些提及错误可能性的会话语境，在那些不提及任何未变成现实的错误可能性的会话语境中，我们似乎更加愿意说某人知道某事。关于我们作出知识归属的意愿和在某个给定的会话语境中什么样的可能性被弄成是显著的这两者之间的关系，詹妮弗·内格尔（Jennifer Nagel）（2010年）最近提出了一个富有争议的心理学解释：我们被要求去考虑这样一个例子，在这个例子中，某人正要作出一个判断，而且我们被告知她的判断可能是错误的，这时我们将我们

的担忧视为她的担忧，并且因为她没有对这些担忧作出恰当的反应，我们就（错误地）对她进行惩罚。内格尔的解释有能力改变当前知识论中的当下这些争论的局面，她的解释揭示了知识的归属敏感于会话语境，这种敏感源自一种已知的认知偏见，这种偏见被称为知识的咒语。现在我们有很好的证据证明内格尔是对的；最近的一系列研究为接受如下观点提供了很好的理由，即不仅我们作出知识归属的意愿敏感于在一个给定的会话语境中什么样的可能性被搞成是显著的，而且这种敏感性最好用认知的自我中心主义来解释（亚历山大等人，即将出版）。

敬请关注

这只是正在发生的事情的一小部分。在过去的一年中，出现了一些很不错的有关直觉多样性的新工作，尤其一些探索以性别和文化为基础的直觉差异的重要的新工作（比如，可见内格尔，2012 年；内格尔，2013 年；斯蒂克［Stich］，2013 年；图利，2013 年），以及实验美学中关于被展示一些糟糕的艺术作品是否会影响我们关于该艺术作品的品质的判断的新工作（梅斯金［Meskin］等人，2013 年），以及实验形而上学中关于决定论和其他方案之间的关系的新工作（毛利［Murray］和纳米亚斯［Nahmias］，即将出版；洛斯和尼科尔斯［Nichols］，即将出版），以及实验的心灵哲学中关于主体经验和现象意识的大众心理的新工作（布克瓦尔特［Buckwalter］和费兰［Phelan］，即将出版）。敬请关注更加令人兴奋的进展吧！

特 别 感 谢

我十分感谢楼巍把本书的英文翻译成汉语，感谢他为了与中国的读者分享这本著作而付出的努力。我还要感谢普莱特出版社(Polity Press)的萨拉·多布森(Sarah Dobson)和上海译文出版社的王巧贞，为她们出版这个中文版所付出的努力。我希望读者们会喜欢这本书，并且发现它既是难读的，又是值得一读的。

参 考 文 献

Alexander, J., Gonnerman, C., and Waterman, J. 即将出版："Salience and Epistemic Egocentrism." 载 J. Beebe (ed.), *Advances in Experimental Epistemology*. New York: Continuum Books。

Buckwalter, W. and Phelan, M. 即将出版："Function and Feeling Machines: A Defense of the Philosophical Conception of Subjective Experience." *Philosophical Studies*。

Buckwalter, W., Rose, D., and Turri. 即将出版："Belief Through Thick and Thin." *Nous*。

Lackey, J. 2007: "Norms of Assertion." *Nous* 41, pp. 594 - 626.

Meskin, A., Moore, M., Phelan, M., and Kiernan, M. 2013. "Mere Exposure to Bad Art." *British Journal of Aesthetics* 53, pp. 139 - 164.

Moore, G. E. 1962: *Commonplace Book 1919 - 1953*.

London：Allen & Unwin。

Murray, D. , and Nahmias, E. 即将出版："Explaining Away Incompatibilist Intuitions." *Philosophy and Phenomenological Research*。

Myers-Schulz, B. , and Schwitzgebel, E. 即将出版："Knowing that Pwithout Believing that P. "*Nous*。

Nagel, J. 2010："Knowledge Ascriptions and the Psychological Consequences of Thinking about Error." *The Philosophical Quarterly* 60, pp. 286 - 306.

Nagel, J. 2012："Intuitions and Experiments：A Defense of the Case Method in Epistemology." *Philosophy and Phenomenological Research* 85：pp. 495 - 527.

Nagel, J. 2013："Defending the Evidential Value of Epistemic Intuitions：A Reply to Stich." *Philosophy and Phenomenological Research* 87, pp. 179 - 199.

Rose, D. , and Nichols, D. 即将出版："The Lessons of Bypassing. "*Review of Philosophy and Psychology*。

Rose, D. , and Schaffer, J. 即将出版："Knowledge Entails Dispositional Belief. "*Philosophical Studies*。

Stich, S. 2013："Do Different Groups Have Different Epistemic Intuitions? A Reply to Jennifer Nagel. "*Philosophy and Phenomenological Research* 87, pp. 151 - 178.

Turri, J. 2013："A Conspicuous Art：Putting Gettier to the Test. "*Philosophers' Imprint* 13, pp. 1 - 16.

Turri, J. 即将出版："The Test of Truth：An Experimental Investigation of the Norm of Assertion." *Cognition*。

致　谢

没有什么东西是独自写出来的，我受益于很多朋友和同事的批评和建议。在他们中间，有一些人需要特别感谢：约翰·亚历山大（John Alexander）、卡梅隆·布克纳（Cameron Buckner）、亚当·阿里科（Adam Arico）、查德·贡纳曼（Chad Gonnerman）、乔纳森·市川（Jonathan Ichikawa）、约书亚·诺布（Joshua Knobe）、S·马修·里奥（S. Matthew Liao）、罗纳德·马伦（Ronald Mallon）、珍妮弗·纳都（Jennifer Nado）、肖恩·尼科尔斯（Shaun Nichols）、马克·费伦（Mark Phelan）、斯蒂芬·斯蒂克（Stephen Stich）和乔纳森·温伯格（Jonathan Weinberg）。在某种程度上，在本书的写作过程中，这些人都有一种以某种重要的方式塑造着本书的洞见；他们中间的很多人还有很多这样的洞见——或许他们还不知道这一点。他们对本书的贡献是无法估量的。

我还要感谢我绝佳的编辑艾玛·胡奇森（Emma Hutchinson）。感谢她的团队的成员，以及普莱特出版社（Polity Press）的两个匿名评审，还有布兰登·尼尔森（Brandon Nelson），他提供了封面上的图画。不管是在风格还是在内容上，由于他们给出的建议，本书都变得更好了。

第五章包含我的论文“实验哲学在哲学上是否相关？”（*Philosophical Psychology* 23，pp. 377 - 389）的很大一部分。我要感谢重印的许可。任何一本书的真正代价是时间；那些花

在写书上的时间,这些时间也可以用来陪伴家人和朋友。若没有我的朋友和家人的爱和支持,这本书将会是不可能的。我希望,尤其是在他们的眼里,这本书值得付出这些代价。

目　　录

001　中文版序言
001　致谢

001　引言
013　第一章　哲学直觉
037　第二章　实验哲学和哲学分析
067　第三章　实验哲学和心灵哲学
095　第四章　实验哲学和哲学方法论
121　第五章　为实验哲学辩护
149　后记

155　参考文献
171　索引
185　译后记

引　言

1. 实验哲学：背景介绍

我们让哲学直觉——*我们会说点什么或事物对我们而言似* 1
乎是怎么样的——为我们做很多的事情。我们提出一些理论，其依据就是它们解释我们的哲学直觉的能力，我们为这些理论的真辩护，其依据就是它们与我们的哲学直觉的总体一致，我们为我们的哲学信念辩护，其依据是它们符合我们的哲学直觉。也许这并不是我们所做的全部事情，也许并不是我们所有人都这样做。但是，在我们当中已有足够多的人这样做了，而且经常是人数众多，以至于至少在某些圈子中思考哲学的这种方式已经成为了思考哲学的*特定*方式。

根据这种思考哲学的方式，哲学家对研究人们的哲学直觉感兴趣，应该是自然而然的。就传统看来，这种兴趣采用了这样一种形式，即内省地考察我们自身的哲学直觉。除了假定我们自身的哲学直觉具有恰当的典型性以外，我们不再需要其他什么东西了。这个方法的问题在于，我们有假定我们自己的哲学直觉具有恰如其分的代表性的习惯，结果发现这是一个*坏*习惯。它忽视了我们的人性的这样一种倾向，即过高地估计了他人认同我们的程度（比如，见菲尔德兹[Fields]和舒曼[Schuman]， 2
1976 年，以及罗斯[Ross]，1977 年），它没能认识到哲学家构成

了一个相当特殊的团体，这个团体不仅由共同的教育背景决定，还由对某些种类的问题以及处理那些问题的方法的共同兴趣所决定。我们需要一个更好的方法。

近些年来，作为一个研究人们的哲学直觉的令人振奋的新方法，**实验哲学**出现了。实验哲学家使用社会科学和认知科学的方法来研究哲学认识，因为比起内省方法，这些方法更适合于研究人们，尤其是**其他**人实际上是怎么想的。这些方法不仅为我们提供了把握相关直觉的更好路径，还可以为我们提供一种洞察力，让我们能洞悉这样一些心理机制（psychological mechanism），①这些心理机制掌管着这些直觉以及这些直觉所有证据上的恰当性。这样一来，实验哲学既可以作为研究哲学问题的更为传统的方法的补充，还可以帮助确定改革那种方法的办法。

实验哲学是一个多元化的运动，在本书中，我们将关注它的三个不同的方案：

1. **实验哲学和哲学分析**：哲学直觉应该能够告诉我们一些关于世界和我们自身的东西，一般说来，正是这些东西让哲学家对人们的哲学直觉感兴趣。这个看法通常是与**哲学分析**联系在一起的。哲学分析常常涉及这样一些论证，这些论证要从关于人们的哲学直觉的断言过渡到关于某个具体哲学理论的真或合理性的断言。有时候，这些理论是与我们关于某物的概念有关的；有时候，它们则与事物本身有关。不管是哪一种，这里的看法都是：哲学直觉应该能够帮助我们回答某些种类的哲学问题，而且一些实验哲学家认为他们自己正在为这个方案作出重

① 这里的“机制”应该被理解为大脑的物理性的机制，而不是一个抽象的概念。——译者

要贡献。这些实验哲学家和更为传统的分析哲学家一样，一般都认为哲学直觉为我们提供了重要的哲学洞见，但是他们相信 3
我们应该采用一些更为适合于仔细研究哲学直觉的办法，换言之也就是社会科学和认知科学的办法。

2. **实验哲学和心灵哲学：**一些实验哲学家认为，关于我们自身，哲学直觉能够告诉我们的东西并不局限于我们关于事物的个人的或共有的概念，他们提出，我们的哲学直觉还可以为我们提供关于我们的心灵之本性的重要的哲学洞见。比起确定哲学概念的准确意义，这些实验哲学家更感兴趣于确定那些影响我们对这些概念的使用的因素。我们的目的不是用我们的哲学直觉来帮助解释我们的哲学概念的意义，而是解释我们的直觉本身，并且借此来揭示某些具有哲学上的重要性的东西，这些东西是与我们的心灵的工作方式以及我们通常理解这个世界的方式有关的。

3. **实验哲学和哲学方法论：**事情是这样的，关于我们的心灵的工作方式和我们思考哲学问题的方式，我们知道得更多，就能产生关于我们在搞哲学时所使用的方法的重要洞见，并且促使一些实验哲学家开始关心直觉在哲学实践中扮演的角色。人们发现，不同的人有着不同的直觉，人们的直觉敏感于很多东西（比如种族、性别、情感及呈现顺序①等），我们既无法预料它们，或许也不希望它们出现。让这个情况变得更加糟糕的是，因为对于所涉及的认知过程我们所知甚少，所以我们现在没有确定该相信**谁的**直觉以及**在什么时候**相信这些直觉的办法。毕竟，证据多样性要求有一些裁定的方法，无法预料的证据敏感性要

① 呈现顺序，原文为“presentation order”，指的是在接受调查问卷时一些问题或场景出现的先后顺序。——译者

求有一些预见的方法，那些不受欢迎的证据敏感性则要求有一些预防的方法。因此，这里所需要的是更好地理解与我们的哲学直觉的产生有关的那些认知过程。借助于更好地理解直觉是什么，它们来自哪里，以及它们受哪些因素的影响，我们可以更好地理解它们在哲学实践中可能扮演的角色。

4 还有其他一些划分这块领地的办法，而且实验哲学中的一些工作不能被很好地纳入这三个范畴，但是这张地形图将足以成为领略实验哲学运动的一个有用的（如果它还不是完全精确与全面的话）向导。无论如何，它将是我们的向导。①

2. 关于范围和覆盖面的说明

一般说来，我们不仅要用书本所涵盖的东西，还要用未涵盖的东西，来衡量那些类似本书的书籍。尽管仍很年轻，但是实验哲学运动已经产生了很多东西，这些东西无法用一个简短的引言来充分地讨论，因此，接下来的介绍会有一些遗漏之过。这些东西中的一些应该被专门提及，应该得到关注，不管它们有多么简短。

意识的实验哲学

近年来，越来越多的人对意识的实验哲学，尤其是对**现象意识**（phenomenal consciousness）（经验）和**非现象意识**（non-phenomenal consciousness）（主体[agency]）的区分产生了浓厚的兴趣。在这些兴趣中，有很大一部分着力于我们对如下几个

---

① 对划分这一领地的其他方式的其他讨论，可见亚历山大和温伯格（2007年），诺布（2007年a），纳德尔霍弗和纳米亚斯（2007年），里奥（2008年）和亚历山大等人（2010年a）。

问题的**大众心理**的理解：为了让某物被认为能够具有现象意识(比如能够看见红色或者感到痛)，需要什么东西；这些所需的东西是否不同于为了让某物被认为能够具有非现象意识(比如能够形成信念、欲望和意图)而需要的东西。初步的发现是很有趣的，这些发现不仅揭示了大众对现象意识的归属①和对非现象意识的归属之间的一种相当复杂的关系(盖瑞[Gray]等人，2007 年；诺布和普林茨[Prinz]，2008 年；阿里科[Arico]，2010 年；胡伯纳[Huebner]等人，2010 年；斯特斯马[Sytsma]和马舍雷[Machery]，2010 年；阿里科等人，2011 年)，还揭示了大众心理对现象意识的归属是与对道德状态的归属联系在一起的(罗宾斯[Robbins]和杰克[Jack]，2006 年；盖瑞和温格纳[Wegner]，2009 年)。由于一幅关于意识的大众心理画卷已经开始产生，在接下来的几年中，将需要更多的工作来解释大众心 5
理对现象意识的归属和非现象意识的归属之间的关系的准确性质，并且还要解释支撑着这两种大众心理的精神状态归属的一种或一些认知过程。

道德哲学和道德心理学

很难[不去]②过分强调那些围绕道德心理学中最近的经验

① 归属，原文为“attribution”，这是一个勉为其难的翻译，它的典型用法是某人将某种东西归之于另外的某人，但其名词形式只能勉强译为“归属”了。就本文而言，下文中的例子有把“具有知识”归属于某人，把“道德责任”归属于某人，等等。——译者

② 参照上下文的意思，按照中国人的理解习惯，这句话的意思本应是“由于这些兴奋和争议实在太多、太重要了，以至于要做到如下这一点是很困难的：不去过分地或者过多地强调那些围绕道德心理学中最近的经验研究而产生的兴奋和争议”，但询问作者后，按照西方人的理解方式，这里的意思应该是“由于这些兴奋和争议实在太多、太重要了，再怎么强调都不为过，所以‘过分地或过多地强调’是不可能的，是很难的”。——译者

研究而产生的兴奋和争议。为了详细地审查我们是如何看待道德和伦理问题的，这些研究既利用了社会科学和认知科学的经验方法，又利用了传统的哲学的概念方法。结果是美妙的，它不仅质疑了处理这些问题的一些更为传统的方法，还挑战了哲学伦理学中的一些正统观点。

性格和情境

道德哲学家总是认为，人们具有坚实的性格特征，这些特征决定了我们的大部分道德和伦理行为。我们的性格并不总是好的，且会受到背景因素的影响，然而正是我们的道德品性解释了我们的道德和伦理行为的全部模式。这个观点在道德哲学家中甚是流行，近些年来，由于道德心理学中的经验研究的影响，这一观点遭到了质疑。结果是，人们在不同的情境下有着不同的行为，哪怕是微小的背景差别，都会影响我们的道德和伦理行为。（比如，可见艾森[Isen]和列文[Levin]，1972 年；达利[Darley]和巴特森[Batson]，1973 年；以及米尔格兰[Milgram]，1974 年）约翰·多丽斯（John Doris）（1998 年和 2002 年）和吉尔伯特·哈曼（Gilbert Harman）（1999 年）提出经验结果已经为基于性格的德性伦理学带来了一些难题，他们转而提出一种被他们称为**情境主义**（situationism）的观点，根据这一看法，我们的道德和伦理行为并不由坚实的性格特征所决定，而是很大程度上由背景因素所决定。（最近对情境主义质疑的讨论，包括为基于性格的德性伦理学的辩护，可见库普曼[Kupperman]，2001 年；斯里尼瓦萨[Sreenivasan]，2002 年；安纳斯[Annas]，2003 年；坎特卡[Kamtekar]，2004 年；阿皮亚[Appiah]，2008 年；马舍雷，2010 年。）

理性和情感

关于情绪在道德判断中扮演的角色，哲学家已经争论了很

久。最近，这场争论有了一种经验转向。来自神经科学、社会和发展心理学以及精神变态学的证据似乎都指向了这样一个观点，即对道德而言情绪是本质性的。结果人们恢复了对如下观点的兴趣：道德判断以某种方式受到情绪或情感(sentiments)(在某种情况下具有某些情绪的性状)的影响。比如，肖恩·尼科尔斯(2004 年 b)提出，道德判断来自我们对于某些事件的情感性反应和确定错误行为的规范性理论的相互作用。尼科尔斯承认，道德认知的这两个组成部分都是必要的，因为规范性理论阻止了某些行为，我们并不认为这些行为在道德上是错误的(比如礼节规范)，而且我们可以对某些事件有强烈的情感反应，而不必认为它们违反了道德(比如一个孩子摔倒了，擦破了膝盖)。道德认知的这两个部分是截然不同的，它们在**核心道德判断**(即那些涉及阻止了某些行为的"情感规则"的判断，这些行为可以独立地诱发一些强烈的负面情感反应)中走到了一起。

关于情绪和道德判断之间的关系，杰西·普林茨(Jesse Prinz)(2007 年)给出了一个更强的观点，他称之为**强情绪主义**。按照这一观点，我们的道德概念和道德属性(moral property)自身都与我们的情绪反应有着本质的联系。在尼科尔斯认为我们的情绪反应和道德判断之间存在着一种因果关系的地方，普林茨认为这种关系是**构建性的**。(实际上，普林茨认为道德情感既构成了道德判断又构成了道德属性。这给出了这样一种可能性，即某些道德判断可能并不伴随着实时的情绪反应。)

因此，在道德判断与情绪或情感之间似乎存在着一种强因果的或建构性的关系。这给长久以来盛行于道德哲学中的理性在道德判断中扮演着某种重要角色这一看法带来了压力。在最低限度上，我们好像必须在道德列表上为情绪和情感留两个位置。但是，理性的前景可能甚至还没这么好。乔纳森·海德

(Jonathan Haidt)(2001 年)提出,情绪和情感在道德判断中扮7 演着一个如此重要的角色,以至于审慎的理性之所以出现,只是为了在事后对那些道德判断进行**事后的**理性思考。如果这是对的,那么理性就仅仅成为了道德判断的辩护士。

神经伦理学

近些年来,道德心理学的最为有趣的进展之一,就是利用功能神经成像技术(functional neuroimaging technology)来研究道德认知。功能磁共振成像(functional magnetic resonance imaging,简称为 fMRI)技术使得科学家们能够检查在人们作出道德判断时人脑的哪一块在活动,并且借此来更好地理解作出这些判断所涉及的那些不同的认知过程。约书亚·格林(Joshua Greene)用 fMRI 技术提出了一种道德认知理论,他称之为道德的“双进程”理论,按照这一理论,那些典型的义务论判断是与情绪的神经进程联系在一起的,而那些典型的后果主义判断则是与自觉的认知过程联系在一起的(见格林等人,2001 年,以及格林,2008 年)。双进程理论的一个优点就是它有助于我们理解**电车难题**,也就是为什么在某些电车例子中我们会认为牺牲一个旁观者的生命在道德上是被允许的,而在另一些案例中则不是这样的。格林(2003 年,2008 年)和彼得·辛格(Peter Singer)(2005 年)同样利用了来自道德认知的神经科学研究证据,以此来作为关于在道德哲学中该采用哪种道德判断的规范性论证的一部分。他们认为,我们有一些好理由来质疑典型的义务论判断。因为义务论判断可能敏感于一些道德上不相关的因素,并且是和这样一些认知过程联系在一起的,从进化的角度来看,这些认知过程是不会被选择用来追踪(track)道德事实的。尽管这个论证被证明是相当有争议的(比如,可见博克[Berker],2009 年),然而,神经科学不仅可以在道德心理学中,

还可以在道德哲学中扮演某种角色这一看法，在科学与哲学中得到了越来越多的接受。

道德思索和道德行为

当被问及哲学教育的价值时，哲学家们总是首先会提出：仔细思考规范性问题可以对道德主体产生正面的影响。认为仔细的道德思索能够改善道德行为，而这是道德教育的真正的价值。这是一个很好的说法，但是我们有理由担心它可能不是真 8
的。如果道德教育能改善道德上的表现，那么我们可以预计，职业伦理学家将成为道德主体的楷模，并且比起那些不是伦理学家的人，他们有着更好的表现。毕竟，他们把自己的职业生涯献给了道德思索和道德教育，或许他们还深深地关切着道德本身。然而，比起那些不是伦理学家的人，职业伦理学家好像并不更为慷慨（史怀泽格贝尔［Schwitzgebel］和鲁斯特［Rust］，即将出版），在政治上也并不更加尽心尽责，也并不更有礼貌（史怀泽格贝尔等人，即将出版），甚至更有可能作出某些在道德上令人反感的行为（史怀泽格贝尔，2009 年）。这使得道德教育在本质上是改善性的这一观点很难继续维持下去了，这还表明道德教育的价值并不在此。（有趣的是，经验性研究不仅表明了职业伦理学家并不是道德主体的楷模，还揭示了哲学家也许在某种程度上已经知道了这一点。可见史怀泽格贝尔和鲁斯特，2009 年。这使我们很难解释本段前面提及的那个说法最初是如何出现的。）

科学的实验哲学

最后有必要简要地提一下科学的实验哲学中的一些非常有趣的工作，这些工作旨在理解我们是如何理解一些重要科学概念的，比如**解释**和**基因**，如果就举两个例子的话。

解释

我们想要理解这个世界，这至少部分地意味着我们能够**解释**为什么一些事情发生而另一些没有发生。但是，要解释为什么某事发生(或没有发生)是什么意思呢？为什么能够做到这一点对我们而言如此重要呢？塔尼亚·罗布伦佐(Tania Lombrozo)最近的经验研究(2006 年，2011 年，以及罗布伦佐和凯瑞[Carey]，2006 年)逐渐搞清了这两个问题。结果发现，我们似乎更加青睐于功用性解释(用理由、功能和目的来解释)，由
9 于我们倾向于认为在给出解释的能力中存在着一些本质上很有价值的东西，这种感觉最好用我们的日常生活与科学生活中解释所具有的那种工具性价值来解释。能够解释为什么某事会发生，这一点支持着我们更为实际的目的——预测和控制，而那些自身对实际好处负责的解释的性质又能对我们产生影响。我们因那些性质而重视解释，恰恰是因为这些性质带来了实际的好处，即使我们并不是自觉地意识到是这些性质对那些实际的好处负责的。

基因

按照概念的“经典理论”，概念是与一组特征联系在一起的。就单个而言，这些特征是必要的，就总体而言，它们是充分的。一些科学概念似乎符合这一模型，而另一些则不是这样的。(当然非科学概念也是一样的。)一些科学概念是模糊的，或者是含义不明的，它们在各种科学背景中得到使用，并有着特定的意义。保罗·格里菲斯(Paul Griffiths)和科罗拉·斯托兹(Karola Stotz)(2006 年，2007 年)提出**基因**就是这样的概念；“某个东西是一个基因”的意思至少部分地取决于这个概念在其中被使用的那个特定的科学背景。但是，如果**基因**这个概念是混杂的，如果它的意义随着背景的改变而改变，那么能够确定哪

些不同的意义在起作用以及一个特定意义在何时起作用，就变得重要起来了。结果发现这并非易事。它不仅需要对**基因**概念在其中被使用的各种科学背景，还需要对这个概念在生物学历史中演变的方式进行仔细的经验研究。我们发现，这是一个其意义随着科学背景的变化而变化的突出典型。

3. 一个被置之不理的问题

在我们开始之前，有必要简要地说一下实验哲学到底是不是哲学。① 说它不是哲学，说它只是实验心理学或哲学的心理学，这种指控变得时髦了起来。通常说来，这一指控出现的时候还带着一层包装，即一套关于何种问题才是真正的哲学问题以 10
及在试图回答这些问题时可以采用哪些方法的说辞。以这样的方式来对这一指控进行包装，其问题就在于它带有哲学帝国主义的味道。除非我们采用一套完全狭隘并且貌似独断的标准，否则我们很难坚持认为实验哲学的问题并不是真正的哲学问题，或者社会科学和认知科学的方法并不是合法的哲学方法。然而，真正的问题并不在于这种指控以什么面貌出现，而在于这种指控本身。为了说清楚这一点，让我们假设实验哲学**就是**实验心理学。那又怎么样呢？除非我们对理智工作进行一种严格的、不允许有重叠的划分，否则，如下这一事实并不意味着它不是在用恰当的哲学方法处理真正的哲学问题：实验哲学所提出的问题是与人类的认知有关的，而这些问题总是与实验心理学联系在一起的，同时实验心理学使用了社会科学和认知科学的

① 对这一问题的讨论，还可见纳德尔霍弗和纳米亚斯（2007 年），诺布（2007 年 a，2007 年 b），诺布和尼科尔斯（2008 年），以及索默斯（Sommers）（2011 年）。

方法。这些问题可以具有心理学的和哲学的重要性,同时一种方法可以超越传统学科的界限。如果是这样的话,那又该说什么呢?我怀疑这种指控的动机通常是担心一些真正的哲学问题无法用社会科学和认知科学的方法来回答。没错,但这仅仅意味着不是所有的哲学都是实验哲学。完全正确。接下来就是是实验哲学的那一部分哲学的故事。

第一章
哲 学 直 觉

> 贝勒(George Bealer)这样做。齐硕姆(Roderick 11
> Chisholm)经常这样做。大多数哲学家都公开而不带辩解地这样做,实际上,其余的人也这样做,尽管他们中的一些人可能会否认这一点。他们都在做的事,就是诉诸直觉来建构、塑造、加工他们的哲学观点。
>
> ——希拉里·科恩布利思(Hilary Kornblith)(1998 年)

1. 引　　言

哲学直觉在当代哲学中扮演着一个重要的角色。哲学直觉为我们提供我们的哲学理论需要对其作出解释的材料,提供我们在论述这些理论为真时援引的证据,还提供我们相信这些理论为真时诉诸的理由。这样一来,直觉证据在哲学中的角色和相应的认知地位便类似于知觉证据在科学中的角色和相应的地位。① 因为实验哲学产生于这样一种看待哲学的方式,所以理解实验哲学就需要对哲学直觉及其在当代哲学中扮演的角色进行
缜密而仔细的考察。因此,在一开始,我们要以一个关于哲学直 12
觉及其在当代哲学实践中扮演的角色的故事来作为开头。

2. 直觉和哲学实践

哲学家所关心的是对这样一些理论作出评价，与这些理论有关的是诸如知识、确证、意义、道德责任以及道德上正确的行为(morally right action)，等等。② 知识就是确证了的真信念吗？一个信念是被确证的，这只是因为它是由一个可靠的认知机制引起的吗？一个名称所指的是不是任何一些这样的对象，这些对象独一无二地或最好地满足与这个对象相联的摹状词？是不是只有当某个人本来可以做出其他行为的时候，这个人才能因某项行为而负道德责任？如果同等条件下某个行为给最大多数的人带来了最大的利益，这个行为就是道德上正确的行为吗？当面对这种类型的问题时，哲学家通常诉诸关于真实的或想象的例子的哲学直觉。我们提出一些哲学理论，其依据就是这些理论解释我们的哲学直觉的能力，我们将它们作为这些理

---

① 贝勒(1992 年，1996 年 a，1996 年 b，1998 年)和索萨(1996 年，1998 年)最明确地给出了这一类比。还可见内格尔(Nagel)(2007 年，即将出版)。这种思考问题的方式并未回答如下这一问题，即将直觉(或知觉)视为证据是否会牵涉到将心理状态(或关于心理状态的命题)视为证据或将那些心理状态的**内容**视为证据。正如我们在后面将会看到的那样，并不是所有哲学家都有这样一种看待哲学的方式，即使是那些认为直觉在哲学实践中扮演着某种角色的人。比如，科恩布利思(1998 年，2002 年，2007 年)和威廉姆森(2004 年，2005 年，2007 年)反对直觉的**证据性**角色，但是坚持它们在哲学中仍然扮演着某种角色，金德勒(Gendler)(2007 年)最近也提出，哲学思想实验的作用并不是生产可以用来支持或反对某个哲学观点的证据，而是有助于使哲学理论在认知上易于把握。

② 一些哲学家认为这些理论是在告诉我们一些关于心理世界(换言之，就是关于我们对于事物的个人的或共有的心理概念)的东西，而其他一些人则认为这些观点是在告诉我们一些关于非-心理世界(换言之就是关于事物本身、柏拉图的形式、弗雷格的概念等)的东西。可见戈德曼和普斯特(1998 年)和戈德曼(2007 年)对此的讨论。我们将在第二章中重拾这一区分，在此之前我们将置之不理。

论为真的证据，作为相信这些理论的理由。

为了更好地搞清楚这是如何起作用的，让我们来看一些例子吧。很久以来，哲学家认为知识就是确证了的真信念。现在他们不这样认为了。葛梯尔(Edmund Gettier)(1963年)给出了两个假定的例子，这两个例子说的是一个人从一个确证了的*假*信念 p 中推论出一个真信念 q，并且以此为基础形成了一个确证了的真信念 q，而 q *似乎*①并不算作是知识，以此，葛梯尔改变了我们的看法。② 葛梯尔的一个例子是：

> 假设史密斯和琼斯都申请了某个工作。并且假设对于下面这个合取命题，史密斯有着充足的证据：
>
> (d) 琼斯是将要获得那个工作的人，并且琼斯的口袋里有十个硬币。
>
> 史密斯对于(d)的证据可能是公司总裁曾经让他确信最后琼斯将会被选中，而且他，史密斯，十分钟之前曾经数过琼斯口袋里的硬币。命题(d)蕴含着：
>
> (e) 那个将要得到工作的人口袋里有十个硬币。
>
> 让我们假设史密斯看到了从(d)到(e)的推断，并且以 13
> (d)为依据而接受了(e)，对于(d)他有充足的证据。这样，史密斯相信(e)是真的，这很明显是确证了的。
>
> 但是，再设想一下，史密斯不知道是他自己而不是琼斯获得了工作。而且，史密斯同样不知道他自己的口袋里也

---

① 似乎，原文为"seem"，作者用这个词强调的是人们关于某个命题、理论或信念的"直觉"，下文中的"对我们而言"("seem to us")也是一样。——译者

② 并不是说*每个人*都改变了他或她的看法。比如，威瑟森(Weatherson)(2003年)提出我们相信知识是确证了的真信念的理由压倒了由葛梯尔那两个假定的例子所引发的哲学直觉所提供的证据。

> 有十个硬币。因此命题(e)是真的,尽管命题(d)——史密斯正是从这里推断出(e)——是假的。在我们的例子中,下面这些都是真的:(1)(e)是真的;(2)史密斯相信(e)是真的;(3)史密斯相信(e)是真的,这一点是确证了的。(葛梯尔,1963年,p. 122)

史密斯有一个确证了的真信念,即那个将要得到工作的人口袋里有十个硬币。然而,**似乎**他并不知道这一点。相反,应该清楚的是,史密斯并**不**知道那个将要得到工作的人口袋里有十个硬币,即使他的信念既是确证的,又是真的。(人们通常认为这里的错误就在于史密斯没有通达恰当的真之制造者[truth-maker]的道路——或者至少使他的信念被确证的东西并不是使该信念为真的东西——因此,他的信念为真,这一点纯属运气。)我们理应**看到**这一点,而这一点理应成为反驳"只要一个人的真信念是确证了的,那么这个人就知道p"这样一个断言的充分证据。① 如果史密斯确实相信那个将要得到工作的人口袋里有十个硬币这一点是确证了的,同时史密斯却不知道这是真的,那么知识就不能仅仅是确证了的真信念。

在关于认知确证的性质的讨论中,哲学直觉扮演着类似的角色。比如,考虑一下如下这个观点:某人的信念p被认为是确证的,只有当该信念是由一个可靠的认知过程所引起,或者在因果上得到这个认知过程的支持。虽然一些哲学家发现这个观点很有吸引力,但是当今大多数哲学家都赞成由一个可靠的认

---

① 重要的是要记住当哲学家诉诸直觉并将其当作证据时,他们不是(至少一般说来不是)只诉诸自己的直觉。哲学的分析并不是纯粹的思想自传,其理由很充分:在论辩上,诉诸某人自己的直觉是无效的。进一步讨论可见亚历山大和温伯格(2007年)。

知过程引起或在因果上得到这个认知过程的支持，尚不足以使某人的信念 p 成为确证的信念。为了搞清原因，考虑下面这个假定的例子：

> 假设有这样一个人，我们称之为特鲁特普先生，[①]他接受了一次脑部手术，实施者是一个搞实验的外科医生，此人发明了一个小装置，这个装置既是一个非常灵敏的温度计，又是一个能产生思想的计算设备。我们可以称这个装置为
> “温度电脑”。[②] 这个装置被植在了特鲁特普的脑子里，而 14
> 且这个装置的末端（并不比针头大）悄悄地露在了他的头皮外，并且起到一个把关于温度的信息传递到大脑内计算系统中的感应器的作用。然后，这个装置向他的大脑发送一个信息，使他想到外在感应器记录下来的温度。假设这个温度电脑是很可靠的，因此他的想法就是关于温度的正确想法。总之这是一个可靠的信念形成过程。最后，我们设想一下，他并不知道温度电脑已经被放在了他的脑子里，他只是有点奇怪为什么他总是会不由自主地想到温度，而且他从未用温度计来确定这些关于温度的想法是否正确。他不假思索地接受了这些想法，这是温度电脑的另一个功能。以这样的方式，现在他认为，并且接受温度为[华氏]104度。而现在的温度确实是 104 度。那么，他知道现在温度是 104 度吗？（莱勒，1990 年，第 163—164 页）

---

① 原文为 Mr. Truetemp，还可以意译成“真实温度”先生。——译者

② 原文为 tempucomp，应该是将温度（temperature）和电脑（computer）结合了起来。——译者

特鲁特普的温度信念是由一个可靠的认知过程引起的，然而似乎他并不知道现在温度是 104 度。事实上，应该清楚的是特鲁特普并不知道现在是 104 度，而且这本应该给“一个人的信念 p 是确证了的，如果该信念是由一个可靠的认知过程引起的”这样一个观念带来麻烦。如果特鲁特普不知道现在是 104 度，即使他的“现在是 104 度”这个真信念是由一个可靠的认知过程引起的，那么很明显要确证我们的信念，还需要一些超出它们是由一个可靠的认知过程引起的这样一个简单事实的东西。它们由可靠的认知过程引起，这可能是必要的，但是很明显这并不是充分的。

我们已经查看了两个直觉在当代知识论中扮演着某种角色的例子。当然，直觉在当代哲学中的作用并不局限于知识论。让我们考虑一个被广泛讨论的语言哲学的例子。根据指称的摹状词理论(descriptivist theory of reference)，每一个专名都有一个摹状词与之相关联。这个摹状词指定了一组性质，如果某个单个对象独一无二地或最好地满足了这个摹状词，换言之就是它具有所有与摹状词相关联的性质，或者，比起其他任何对象，它具有更多这样的性质，那么它就是这个给定的专名的所指；相反，根据指称的因果理论(causal theory of reference)，一个专名被引入一个语言共同体，是为了指称一个特定的个体。因此，只
15 要借助一种连续使用的因果顺序，它的用法仍然与那个个体相关联着，换言之就是只要某人从他人用这个名称来指称那个单个对象那里学会了用这个名称来指称那个单个对象，并且以此类推，一直回到这个名称最初被引入来指称那个单个对象的那个点上，那么，这个名称就将继续指称这个个体。

哪种观点是对的呢？在 20 世纪 70 年代之前，大部分分析哲学家都支持摹状词理论，他们相信名称所指的这样一些对象，

这些对象独一无二地或最好地满足与这些对象相关联的摹状词。随着《命名与必然性》一书(克里普克,1980年)的出版,事情(至少在一段时间内)起了变化。[①] 在书中,索尔·克里普克(Saul Kripke)给出了一系列强有力的论证来反对摹状词理论。在这些论证中,最著名的一个论证是以对某个人的描述开始的,这个人说哥德尔是那个证明了算术不完备性的人。此人对算术很是了解,足以对不完备定理进行独立的阐述,但是关于哥德尔他所知道的全部东西就是哥德尔是证明算术不完备性的那个人。带着这个想法,克里普克邀请读者来考虑一下如下这个假定的例子:

> 假设哥德尔实际上并不是这一定理的作者。实际上,是一个被称为"施密特"的人做了这项工作,在多年前的维也纳,他的尸体被离奇地发现了。他的朋友哥德尔以某种方式取得了这份手稿,此后这份手稿就被认为是哥德尔的了。(第83—84页)

根据指称的摹状词理论,当我们这个虚构的主角说哥德尔是证明了算术的不完备性的那个人时,他指的是施密特,因为正是施密特独一无二地满足了与"哥德尔"这个词相联系的那个摹状词(即那个证明了算术的不完备性的人)。然而,似乎他并不是在指施密特。相反,似乎他指的是哥德尔(即使他关于哥德尔的一些信念最后发现是错误的)。而且,这是反对摹状词理论的

---

① 随着二维语义学的兴起,摹状词主义近些年来有所复兴。可见查默斯(Chalmers)(1996年)、杰克逊(Jackson)(1998年a,1998年b)和斯塔纳克(Stalnaker)(1999年)。

一击,即它是与这种直觉相悖的。指称的摹状词理论告诉我们这个人所指的是施密特,而对我们而言似乎很明显这个人指的
16 是哥德尔,这个事实揭示了名称所指的并不是这样一些对象,这些对象独一无二地或最好地满足与这些对象相联系的摹状词。相反,指称的因果理论成功地解释了为什么对我们而言似乎这个人指的是哥德尔:他对这个名称的使用,是与引入这个名称来指称哥德尔恰当地联系在一起的。指称的因果理论能够解释为什么这个人所指的似乎是哥德尔,这揭示了指称是由最初的命名活动固定下来的,而且后面对这个名称的使用之所以能成功,是因为这些使用与最初那个行为因果地联系着。

除知识论和语言哲学外,哲学直觉还可以在当今关于自由意志和道德责任之关系的讨论中扮演重要的角色。比如,很多支持不兼容主义(这种观点认为自由意志、道德责任与因果决定论是不兼容的)的论证依赖于这样一个原则,即只有当一个人可以做出其他行为的时候,她才应该负道德上的责任。如果只有当一个人可以做出其他行为的时候她才应该负道德上的责任这一点是真的,同时,如果因果决定论规定了在过去的事实和自然律被给定的情况下,我们只能做那些我们确实做出的行为,那么因果决定论和道德责任就是不兼容的。但是,这个一般被称为"另一种可能性原则"的原则是真的吗?哈里·法兰克福(Harry Frankfurt)(1969年)给我们提供了担心事情也许并不是这样的理由。法兰克福请读者设想一个名为琼斯的人。基于他自己的理由,这个人决定做某个行为,但是,另一个人用某种惩罚来威胁他(这种惩罚是如此严厉,以至于任何一个有理性的人都会屈服于这一威胁),要求他做那同一个行为。琼斯既不是一个做事不考虑结果的不理性的人,也不是一个"屈服于威胁"的人,以至于会"由于受到威胁,便不顾他已经作出的决定而做

出那个行为”(法兰克福,1969 年,第 832 页)。相反,

> 这个威胁压倒了他,正如它会压倒任何一个有理性的人那样,而且他可能会全然地屈服于这一威胁,如果不是他已经作出了一个决定并且这个决定与那人要求他做出的行为彼此一致的话。不过,实际上他以受威胁前的决定为依据做出了这个行为。当他在行动的时候,事实上他并不受这个威胁所驱使,而仅仅出自那最初的考虑,正是这一考虑让他执行这一行为。并不是这种威胁促使他行动的,尽管倘若他没有为自己提供一个充分的动机来执行这一行为的 17
> 话,那就将会是这种威胁促使他这样行动的了。(法兰克福,1969 年,第 832 页)

根据另一种可能性原则,琼斯无需为他的行为负道德责任,因为他不能做其他的行为。然而,似乎他要为他的行为负道德责任。在受到威胁之前,琼斯决定做这个行为,他是以这一个决定,而不是以随后的威胁为依据而行动的。但是,如果琼斯要为他的行为负道德责任,那么即使一个人不可能做其他行为,他可能也要为其行为负道德责任了,特别是,如果使他不可能做其他行为的处境实际上并不是导致他这样行动的首要东西的话。

让我们考虑另一个例子,这次这个例子来自于规范伦理学。很多道德哲学家都相信,某个有害的行为在道德上是不被允许的,除非这样做是另一个行为的不可避免的副作用(side effect),[1]而这另一个行为的初衷是想要带来某个道德上为善

---

① 原文为“Side effect”,很多人翻译成“边际效应”,就本文的语境来说,这是不够通俗和明确的。——译者

的事态(a morally good state of affairs)。尽管在产生某些道德上为善的事态的过程中引发某些有害的副作用是被允许的，然而，把这同一些有害的作用作为引起那种事态的**手段**，却是不被允许的。一些道德哲学家赞同这一看法(一般被称为“**双重作用学说**”)，其理由是这种看法有助于解释为什么我们对如下两个假定的例子有着不同的反应：

电车

假设你是一辆电车的驾驶员。电车失控了，而前面出现了五个正在修理轨道的轨道工。此时，轨道正要穿过山谷，两边都很陡，因此，如果你不想撞倒那五个人，你必须把车停住。你踩了踩刹车，但是天哪！刹车坏了。现在你突然看到有一条通向右边的岔路。你可以把电车开到岔路上去，并挽救前面直行这条路上的五个人。不幸的是……岔路上也有一个轨道工。这个人和那五个人一样，都不可能及时离开轨道了，因此如果你将电车开到岔路上，你将会杀死这个人。(汤姆森[Thomson]，1985 年，第 1395 页)

移植

18 假设你是一名外科医生，一名真正伟大的外科医生。你的主要工作是移植器官，而且你是一名杰出的外科医生，你移植的器官总能存活。此时，你有五个需要器官移植的病人。两个人各需要一只肺，两个人各需要一个肾，第五个则需要一个心脏。如果他们今天不得到这些器官，他们都将死去；如果你今天为他们找到了这些器官，你可以移植它们，他们就能活下去。但是上哪儿去找这些肺啊肾啊心脏呢？时间就快到了，这时你接到一个报告，说一个年轻人刚

> 刚来到你的诊所做年度体检，这人有着完全一致的血型，而且身体状况极佳。你看，你有了一个可能的捐献者。你只要将他切开，把他的各个器官分给那五个需要它们的人。你问了这个人，但是他说："不好意思。我表示深切的同情，但是不行。"（汤姆森，1985年，第1396页）

双重作用学说解释了为什么改变电车的方向**似乎**在道德上是被允许的，而牺牲一个健康的人则是不被允许的。在电车的例子中，我们并不想要杀死一名工人，并以此作为拯救其他五人的**手段**。他的死亡是这样一个行为的不可避免的副作用，这个行为本想带来一种在道德上为善的事态。相反，我们确实想要杀死这个健康的人来拯救其他五人。他的死亡是我们想要带来一种在道德上为善的事态的手段。在带来道德上为善的事态的过程中带来有害的副作用，这在道德上是被允许的，而用有害的手段来达到道德上为善的目的，这在道德上是不被允许的，规定了这一点，双重作用学说便解释了我们关于这两个例子的直觉，这就是它的长处。①

3. 哲学直觉

类似这样的例子挺多的。实际上，哲学直觉在当代的哲学实践中扮演着一个如此重要的角色，以至于乔治·贝勒(George Bealer)(1992年，1996年a，1998年)说哲学直觉是"我们的标

① 对双重作用学说和其他版本的电车例子——包括对被认为产生了与双重作用学说不一致的直觉证据的电车例子的讨论，可见福特(Foot)(1967年)、汤姆森(1985年)、卡姆(Kamm)(1989年、2007年)和里奥等人(即将出版)。

19 准的辩护过程"的一部分。在描述哲学直觉在为如下观点——即知识不仅仅是确证了的真信念——辩护时所扮演的证据性角色时，贝勒写道(1996 年 a)：

> 从前，有很多人接受知识是确证了的真信念这一信条。但是，今天我们有很好的证据来证明它的反面才是对的，那就是我们的如下直觉：类似葛梯尔文章中所描述的那种情况是可能的，同时处于这种情况中的人们并不知道那些东西。这个例子外加无数其他的例子揭示了，根据我们的标准的辩护过程，直觉被当作证据(或当作理由)而使用了。(第 122 页)

贝勒不仅仅给出了一个描述性的社会学断言，即哲学直觉实际上是我们标准的确证实践的一部分，他还提出了一个规范性断言，即它们*必须*是这样的。这个规范性断言是他对经验主义的不一致性的著名论证的一部分，这一论证的基本观点是：经验主义者无法用任何一种不包括哲学直觉的确证方法来为自身的方法论约定辩护。比起贝勒，其他哲学家要走得更远一些，他们不仅仅提出哲学直觉是我们的标准确证过程的一部分，还提出这一事实是使哲学的方法论具有独特性的原因的一部分。比如，贾内特·列文(Janet Levin)(2005 年)写道：

> 知识是确证了的真信念吗？因果关系只是事件的持续不变的同时发生吗？有意识的经验仅仅是某种神经活动吗？这些问题，就其与事物的本性(而不仅仅是我们的概念的结构)的关系而言，是模态论题，因此无法仅仅用经验方法来研究。相反，或者正如哲学家所一般认为的

那样，要回答这些问题，就得咨询我们的哲学直觉；换言之就是那些有关这样一些可能性的极具说服力的陈述，这些可能性产生于严肃而深思地设想这样的一些情景，这些情景可以作为这些论题的反例……在哲学史上，在那些直觉上令人信服的反例之前，去否定或修改这些论题的过程都是哲学论争的特点，也许还是它的决定性因素。（第193页）

还有，正如戈德曼（Alvin Goldman）（2007年）所言： 20

将哲学方法和科学方法区分开来的东西，就是哲学方法大量而公然地依赖直觉。尤其是哲学家在搞哲学"分析"的时候，他们经常会专注于直觉。为了确定什么是知识、指称、统一性、因果性（或者什么是知识、指称、统一性、因果性的概念）的时候，哲学家一般都会考虑一些实际的和假定的例子，并且追问这些例子是否提供了目标范畴或概念的实例。人们对这些实例的精神反应通常被称为"直觉"，而且这些直觉被当作正确答案的证据。就最低限度而言，它们是这些例子是或者不是知识、指称、因果性等的实例的证据。这样一来，直觉在哲学活动的某个特定部门中扮演着一个尤为关键的角色。（第2页）

但是，什么是哲学直觉呢？答案有很多，而且这些答案经常反映了关于我们该如何回答这些问题的不同态度。一些哲学家鼓励我们去咨询我们关于直觉的直觉；其他一些哲学家则建议我们要特别注意从第一人称视角看来的，从内省上说直觉到底是什么东西；还有其他一些人则建议我们仔细地审查哲学家在

实践中是拿什么来作为直觉证据的。① 正如乔纳森·温伯格和约书亚·亚历山大(即将出版)所指出的,这些不同的方法已经产生了从**薄观念**(thin conceptions)到**厚观念**(thick conceptions)的一系列观念。薄观念认为哲学直觉是某些特殊种类的精神状态(或精神事件),特别是信念或者相信的倾向;厚观念则为何种精神状态才算作真正的哲学直觉添加了一些额外的条件。让我们从这些观念中择取一些,来简要地看一下吧。

信念观念

一些哲学家认为直觉仅仅是信念,或者也许是相信的倾向。② 大卫·刘易斯(David Lewis)(1983 年)采纳了这一看法:

---

① 林奇(2006 年)赞同第一种策略,贝勒(1998 年)、索萨(1998 年)和普斯特(2000 年)赞同第二种,威廉姆森(2007 年)赞同第三种。尽管这里我们不讨论这个问题,但是第三种策略可能是最好的。第一种策略的问题之一在于我们是否拥有关于直觉是什么的**任何**直觉,这一点是不清楚的(或者至少是这样的,直觉到底是不是那种我们对其有某种直觉的东西,这一点将最终取决于直觉会是什么),因此这一策略可能无法在决定哪一个看法正确的时候提供太多的帮助。同时,第二种策略的问题之一在于,在那些诉诸从第一人称视角看来直觉是什么的哲学家当中,也几乎没有关于直觉是什么的一致意见,因此,与其说这一策略能扫清困惑,不如说它将带来困惑。第三种策略避免了哲学直觉是不是那种我们对此有直觉的东西这一棘手的问题,同时也避免了这样一种混乱——这种混乱是由关于对我们而言直觉似乎是什么的各种第一人称的、内省的考察所引起的。它还有另一个优势,那就是理论上的谨慎性,除了把握哲学直觉在哲学实践中的角色的最低需要外,它没有为对哲学直觉的考察添加任何东西。

② 当然,这并不意味着所有的信念或者相信的倾向都是哲学直觉;几乎所有哲学家都想要对哪些种类的信念或者相信的倾向视为哲学直觉加以限制;比如,大多数哲学家都提出哲学直觉应该是**直接的**,或者至少**不是明显推断性的**。在这里需要指出的是,本节提到的哲学家对于哲学直觉是不是信念或相信的倾向这一问题均持中立态度,不过其他那些支持某种信念方法的哲学家对待这一问题就更有偏袒性了。比如,索萨(1998 年)和厄伦巴赫(Erlenbaugh)、莫利努克斯(Molyneux)(2009 年 a,2009 年 b)提出了一种明确将哲学直觉视为相信的倾向或性状(disposition)的看法。

> 我们的“直觉”只是观点而已；我们的哲学理论也是一样。有些是稀松寻常的，有些是精深的；有些是具体的，有些是一般的；有些被坚定地持有，有些则要差一些。但它们都是观点，对一个哲学家来说，一个合理的目标就是将它们带入到某种平衡之中。（第Ⅹ页） 21

彼得·范·因瓦根(Peter van Inwagen)(1977年)对此表示同意：

> 我们的“直觉”就是我们的信念——在某些例子中，或许是使某些信念吸引我们的倾向，这种倾向“促使”我们走到接受某些命题的方向上去，而不是直接使我们接受它们。（哲学家把他们的哲学信念称为直觉，因为比起“信念”，“直觉”听上去更具权威性。）（第309页）

同时，我们可以怀疑范·因瓦根并没有准确把握我们的这样一种动机，即把某些信念作为哲学直觉挑选出来。蒂莫西·威廉姆森(Timothy Williamson)(2007年)最近提出，我们有充分的理由来以如下方式看待它们：更具限定性的观念无法恰当地把握哲学家在实践中作为直觉证据而诉诸的全部东西，同时，更有包容性的观念则可以让普遍的直觉怀疑主义变得越来越没有吸引力，借此把哲学直觉和怀疑主义的质疑所带来的威胁隔绝开来。① 我们仍然有其他

---

① 在这里，有一点可能是值得注意的，那就是威廉姆森实际上提出诉诸哲学直觉作为证据的实践依赖于一种对哲学证据之本性的错误理解，他主张哲学家“最好不要使用‘直觉’这个词及其同源词”，因为我们对直觉证据的强调并不能帮助我们“回答那些有关证据之本性的问题，而以一种貌似提供了答案，实际上却并未这样做的方式回避了它们”(2007年，第220页)。我们将在第五章讨论威廉姆森对哲学直觉的哲学重要性的反驳，还可以见亚历山大(2010年)。

的一些理由来担心信念观念实在太薄了：很多哲学家想说，哲学直觉和其他种类的精神状态之间存在着重要的区分，同时，我们可以担心这种信念观念并没有让我们处在一个能够标明这些区分的位置上。[1]

现象学观念

标明哲学直觉和其他种类的精神状态之间的区别的一种方式，就是提出真正的哲学直觉有着一种特殊的现象学特征，这种特征乃是直觉经验在内省上可被把握的那些方面。哲学直觉给我们一种确定的印象，它或许是主观上令人信服的东西，或许是
22 必然为真的东西。贝勒(1998 年)接受了这一看法。他提出，哲学直觉是*理智上似乎为真的东西*(intellectual seemings)，但是有着*必然性的外表*(appearance of necessity)。[2] 当对我们而言似乎 p 必然为真时，我们有直觉 p。这意味着直觉并不是信念，因为即使对我们而言*似乎*[3]某些命题为假，我们也相信这些命题为真(比如，考虑一下一个无限集可以比另一个无限集大这一观点)，同时即使对我们而言*似乎*某些命题为真，我们也相信这些命题为假(比如，考虑一下每个属性都定义了一个集合，这个集合完全收纳了那些具有

---

① 其他人对这一点的讨论，可见温伯格和亚历山大(即将出版)。

② 贝勒在若干篇文章中为这个观点辩护。可见贝勒(1992 年，1996 年 a，1996 年 b)。哲学直觉将它们的命题内容作为必然的东西呈现出来这一看法还得到了邦久(BonJour)(1992 年，1998 年)、普兰廷加(Plantinga)(1993 年)、索萨(1996 年)、海勒斯(Hales)(2000 年)等人的辩护。索萨(2007 年 b)、塔尔伯特(Talbot)(2009 年)、本森(Bengson)(即将出版)最近提出了一种现象学观点，这一观点并不包含哲学直觉将它们的命题内容作为必然的东西呈现出来这一看法，而是各自提出，真正的哲学直觉是"受到吸引的自觉的意识状态"，是"让我们有一种为真的印象，我们却完全不知道为什么它们是这样的"，或者"给我们一种确定的印象"。

③ 此处的英文原文是"seem to us"，下文中还有"appear to us"，译者都将其翻译为"对我们而言似乎……"，在本文中，凡是"对我们而言似乎……"的，说的都是我们的直觉告诉我们的，而不是与我们的信念相关的东西。——译者

这一属性的事物)。① 但是,至少对于当前的目的来说,更为重要的是它还意味着不是所有的直觉都是平等地产生的,我们可以借助如下事实,即哲学直觉是将它们的命题内容作为必然的东西而呈现出来的,来将哲学直觉和其他种类的精神状态区分开来。

这个观点曾经非常流行,然而它也是极其富有争议的,而大部分的争议都围绕着哲学直觉有着必然性的外表这一看法。② 正如乔·普斯特(Joel Pust)(2000 年)指出的,问题在于有很多这样的命题,当我们不去思考它们的模态状况(modal status)③

---

① 贝勒(1998 年)给出了另外两个认为直觉并非信念的理由。首先,他提出信念是有弹性的,而直觉则不是这样;你可以让某人相信几乎所有的东西,然而要让某些命题似乎是真的(在"似乎"这个词的相关的现象学意义上),* 则似乎是不可能的。其次,他认为对于那种你可以对此具有信念的命题是没有什么重要限制的,而对于那种你可以对此具有直觉的命题则存在这样的限制。我们将关注贝勒的前两个理由,因为总的看来它们似乎是坚持直觉并非信念的更强大的理由。第一,是不是所有的信念都是有弹性的(考虑一下你当前存在这一信念),同时是不是足够的心理压力或折磨不能让任何东西对你而言似乎是真的(在"似乎"这个词的相关的现象学意义上),这都是不清楚的。第二,即使确实存在一些命题,关于它们你可以具有信念而无法具有直觉,这并不必然意味着直觉不是信念——它也许只能揭示出直觉是信念的一个子集。

* 原文为"in the relevant phenomenological sense of 'seems'",直觉通常采用某个东西"似乎(seems to)是怎么样的"或者"对我们而言似乎是怎么样的"这样的形式,这里应当这样来理解直觉与现象学的关系:先不谈直觉是否正确,这种直觉本身就是一种(在意识中)被直接给予的东西。——译者

② 另一个争论点是哲学直觉是一种**独特的**命题态度。比如,索萨(1998 年)和威廉姆森(2007 年)都提出,在相信 p 的有意识的倾向中,他们找不到什么理智上似乎为真的东西 p(intellectual seeming that p)。

③ 关于"模态状况",可见 http://onlinelibrary. wiley. com /doi /10. 1111/ j. 1755 -2567. 2006. tb00955. x /pdf。此为 Sven Ove Hansson 的一篇文章。按照文中的说法,有关模态状况的核心观点是:某个哲学陈述或某一系统的哲学陈述的"模态状况"是认为这个或这些陈述为真的人指派给这个或这些陈述的真的"必然性的类型"。比如,一个陈述可能就逻辑的必然性而言是真的,就物理的必然性而言是真的,在人类可以设想的所有条件下都是真的,在所有当下存在的条件下是真的,等等。在某个陈述的模态状况和这个陈述被运用的领域之间有着密切的联系,如果一个陈述的真带有逻辑必然性(它的模态状况是"在逻辑上为真"),那么在假想世界中(在这些世界中,物理学法则完全不同于我们所知道的那些法则)它也是真的。——译者

时，对我们而言似乎它们可能是必然为真的，然而只是因为我们绝不会停止思考它们的模态状况，所以对我们而言它们绝不会是必然为真的。他的看法是：一个命题对我们而言是不是必然为真的，这不仅仅取决于这个命题的特征，还取决于那些特征是否引起了我们的注意。这就留下了这样一种可能性，即那些在恰当的情景下对我们而言似乎**将会**为真的命题**实际上**却绝不会为真，这只是因为情景永远不会是恰当的。结论是：我们是否具有直觉 p 这一点最后取决于我们当时在注意着什么。至少对于某些哲学家来说，这一点似乎是有些反直觉的。当然，这一问题的最自然的答案将会是削弱“哲学直觉将它们的内容作为必然的东西而呈现出来”这一前提条件，其办法就是提出只有当我们不去思考它们的模态状况时它们才会是这样的。但是，正如迈克尔·林奇(Michael Lynch)(2006 年)和基尔克·路德维希(Kirk Ludwig)(2007 年)所指出的那样，一些哲学直觉从未具有必然性的外表，甚至当我们不去思考它们的模态状况时也是这样的。如果这是对的，那么仅仅修改这个前提，也是没有用
23 的。相反，似乎我们必须完全取消这一前提，并承认哲学直觉无须具有任何一种必然性的外表。然而，这样做的问题在于，它让我们处于这样一个位置上，我们不得不去寻找将哲学直觉和其他种类的精神状态区分开来的某些其他方法了。

语义学观念

为了做到这一点，一些哲学家提出哲学直觉必须有一种特定的命题内容。恩内斯特·索萨(Ernest Sosa)(1998 年)为这个方法提供了一个很好的例子，他认为哲学直觉必须有一种**抽象的**命题内容。至于这到底意味着什么，索萨并没有说，而是提出他将会信赖我们对“一个命题是抽象的”这句话的运用性理解

(working understanding),[①]但是,对于当他谈到抽象命题并不提及任何具体的个体——个体的人、地点或事物——时(它们可以提及单个的属性或关系)的意思是什么,他确实提供了一点线索。这一方法的问题在于它仍然太严格了,还是把某些我们想要视其为哲学直觉的东西排除在外了。[②] 为了看清这一点,我们需要仔细地区分哲学观点和用来评价哲学观点的直觉。哲学观点是抽象的。我们提到知识是确证了的真信念、只有当信念是由可靠的认知机制引起时信念才是确证的、名称指向任何独一无二地或最好地满足与名称相联的摹状词的对象,等等,这些观点并未明确提及任何具体的个体。但是我们用来评价这些观点的哲学直觉总是**具体的**。我们说似乎史密斯并不知道那个将要得到工作的人的口袋里有十个硬币,即使他的信念既是确证的,又是真的,我们说特鲁特普并不知道现在是华氏 104 度,即使他的信念是由可靠的认知机制引起的,我们说"哥德尔"指的是哥德尔,即使施密特最好地满足与这个名称相联的摹状词,等等,这些哲学直觉**确实**提到了个体。问题很清楚,如果哲学直觉必须有抽象的命题内容,那么这些就不是哲学直觉。但是,很明显它们是哲学直觉。

病因学观念

语义学和现象学观念背后有一个共同看法,那就是试图在 24

---

① 相比那种理论性的学习和认知,运用性理解强调在对一条原理或事物的使用中来认识、理解它们。——译者

② 路德维希(2010 年)提出语义学方法仍然太宽容了,它包容了某些我们不愿意视其为哲学直觉的东西。正如路德维希所指出的,一些抽象命题——比如不存在最大的质数——对我们而言是过于复杂的,以至于我们无法形成任何关于它们的直觉。

哲学直觉的内容或显现中找到某些东西，这些东西可以被用来将它们和其他种类的精神状态区分开来。问题在于，有太多的东西被稀里糊涂地抛弃了。这些观点太严格了，在它们试图不让太多的东西进来的同时将太多的东西排除在外了。这一点使得某些哲学家提出让哲学直觉与其他种类的精神状态区分开来的东西与它们来自哪里有关，而不是它们是什么，或者拥有这些直觉是什么样的。基尔克·路德维希(2007 年)提出了这一看法，他认为真正的哲学直觉是仅仅基于**概念能力**(conceptual competence)[①]而作出的判断。在相同的思路下，恩内斯特·索萨(1998 年)提出哲学直觉是一种基于**概念理解**而作出判断的倾向，乔治·贝勒(1998 年)主张哲学直觉是**拥有确定概念**的产物，安提·库皮尼(Antti Kauppinen)(2007 年)提出哲学直觉是相关概念的**合格使用者在充分理想的条件下**(在这里他们的判断**仅仅受到语义方面考虑的影响**)将会说出的东西。

这里的共同点是我们可以通过查看概念能力在具体直觉判断的形成中扮演的角色，来区分哲学直觉和其他种类的精神状态。然而，结果证明这是很难做到的。正如我们在后面的章节中将会看到的那样，要精确界定什么是我们概念能力的一部分什么又是我们的概念表现(conceptual performance)的一部分，这并非易事(亚历山大等人，2010 年 a，2010 年 b)。但是，即便我们不去考虑这个担忧，事实仍然是：我们需要某些方法来确定哪些直觉有着恰当的病因学谱系(etiological pedigree)，同时问题就在于这里所涉及的认知过程通常说来是无意识的(赫德森[Henderson]& 霍根[Horgan]，2000 年、2001 年)，或者是无

① 可见 James Higginbotham, *Conceptual Competence*, Philosophical Issues, 1998, Volume 9。——译者

法被内省所把握的(林奇,2006 年)。我们也许有着某种优先的路径来通达我们自己的心理状态,但是我们却几乎没有那种通达这些心理状态的原因的路径(尼斯比特[Nisbett]& 威尔森[Wilson],2002 年),这已足以成为担心我们很难将概念能力的影响和概念表现的作用区分开来的理由了。

方法论观念

有一点是很重要的,那就是要明白病因学方法的问题在于 25
很难搞清它如何在实践中起作用。所涉及的认知过程的性质使得我们很难区分概念能力和概念表现,同时,它们各自的(尤其对内省而言的)不可通达性使得我们很难看清我们如何能够确定哪一种精神状态可被算作哲学直觉。这一点已经使得某些哲学家提出在哲学直觉和其他种类的精神状态之间存在着一种**方法论的**区分。这里的看法是,把哲学直觉和其他种类的精神状态区分开来的东西并不一定是它们来自何处,而是一旦哲学直觉出现,我们该拿它们来做点什么。安提·库皮尼(2007 年)提倡这一方法,提出哲学直觉是这样一些精神状态,它们作为对传统哲学讨论的反思性参与(reflective participation)的一部分,要受到批判性的审查;它们是被一种**哲学思考**的过程所认可的精神状态。这种考察事物的方式和认为哲学直觉有着某种特殊的命题内容、特殊的现象学特征甚或特殊的病因学背景的思考方式是高度一致的。实际上,库皮尼既支持病因学观念又支持方法论观念,提出哲学直觉是相关概念的合格使用者在充分理想的条件下(在这里他们的判断仅仅受到语义方面考虑的影响)会说出的东西,而这些东西恰好就是那种对传统哲学讨论的反思性参与所产生的精神状态。

比起考察哲学直觉的其他方法,方法论观念的长处在于它

提供了一种实际的方式，让我们能够确定哪种精神状态被算作哲学直觉，而且比起现象学或语义学方法，我们这样做似乎不会更加可能把太多的东西排除在外。正如我们在接下来的章节中将会看到的那样，哲学直觉是传统哲学反思的产物这一观点仍然引发了人们对它们的认知恰当性的担忧。传统的哲学反思的认知价值有一种重要且常被忽略的有限性，这一点科恩布利思说得很好：①

26 > 当我们不再去思考我们先于反思(pre-reflective)的信念到底是不是确证的这一问题时，一些彼此不同的偏见便开始起作用了。相比我们遇到的那些反对我们的信念的证据，我们更好地记住了那些支持我们的信念的证据。相比那些我们在其中犯错的场合，我们更好地记住了那些我们在其中搞对的场合。我们有一种倾向，对支持我们的信念的论证，会加以善意的评判，对反对我们的信念的论证，则苛以高标准。当我们在形成关于这样一些过程(借助这些过程，我们先于反思的信念得以形成)的判断时，我们似乎把通盘看来我们在作出判断时是相当可靠的这一点当作了小前提，因此，我们有一种认为我们的信念是以这样一些证据(这些证据在我们自己看来是具有高度决定性的)为基础的强烈倾向，不管这些信念实际上是不是在这种基础上形成的。作为结果，反思的结果往往只不过是对反思评价之前持有的信念的正式认可。这些反思不是有助于纠正……而是倾向于以这样一种方式行动，即把我们先于反思的信

---

① 对这一问题的其他讨论，可见科恩布利思(2002 年)，温伯格和亚历山大(即将出版)，以及温伯格等人(即将出版)。

> 念进一步牢固在我们那张更大的信念之网中。很多反思过程并不是用来纠正我们的先于反思的信念的，而只是加强了我们对它们的确信；这样一来，我们就变成了更加自我满足的（即使并不更加准确的）认知主体（2010年，第5页）。

反思似乎有时候对我们的认知有着相当负面的影响，在接下来的章节中，我们要更加详细地考察这一点。

4. 结 论

尽管我们并未选定一种对哲学直觉的解释，不过我希望我们能够更好地搞清楚当代哲学目前使用的这些不同的观念，并且搞清楚他们各自的强项和弱项。虽然哲学家对哲学直觉的准确性质持不同的态度，然而对于它们在当代哲学实践中所发挥的作用，还是有广泛的一致意见的，我希望我们能更好地弄清楚 27
这种作用到底应该是什么。我们以哲学理论解释我们的哲学直觉的能力为基础来提出理论，并且诉诸哲学直觉，以此作为确定这些理论为真的证据和相信它们为真的理由。在接下来的章节中，我们将会看到实验哲学是如何从思考哲学的这一方式中产生出来的，同时还要审查实验哲学是如何能够既对这种搞哲学的方式进行补充，又有助于确立它被改造的方式的。

第二章
实验哲学和哲学分析

1. 引　　言

当哲学家想要理解类似知识或道德责任这样的东西的性质 28
时,他们经常构造出一些假定的例子,这些例子是用来揭示我们
关于这些东西的哲学直觉的。在哲学中,这些直觉扮演着一个
重要的角色,它们帮助哲学家决定要提出哪种理论并为这种理
论辩护,并且提供这样做的方法。最一般地说来,这种思考哲学
的方式是与哲学分析联系在一起的,这种分析通常会涉及这样
一些论证,这些论证要从人们的哲学直觉过渡到关于特定哲学
理论的真或表面合理性(plausibility)的主张,一些实验哲学家
认为他们正在为这一哲学方案作出贡献。他们所追问的问题与
更为传统的分析哲学家所追问的问题是一样的,但采用的方法
则更适用于仔细地研究哲学直觉,换言之,他们采用了社会科学
和认知科学的方法。

在本章中,我们要看一下实验哲学家最近的一些工作。这
些工作关注的是哲学直觉在关于自由和道德责任的形而上学讨 29
论中,以及在关于知识或我们的知识认定的意愿是否受到风险
(即个人要为出错而付出的代价)或显著性(即在一个给定的会
话语境中,有哪些其他可能性是相关的)的影响的知识论讨论中

扮演的角色。实验哲学家对相关直觉以及塑造这些直觉的认知过程进行了仔细的经验研究，并以此来为这些讨论作出贡献。结果是令人惊奇的，它们不但揭示了一些与我们的形而上学和知识论的直觉有关的东西，还揭示了一些与这些形而上学与知识论的讨论的本性有关的东西，揭示了哲学直觉目前是否适合于将这些讨论推向前进。

2. 实验哲学、自由和道德责任

关于自由和道德责任的哲学直觉通常以这样一个假设为起点，即人们是**天然的不兼容主义者**，他们相信自由意志和道德责任都与因果决定论不相容。[①] 这个与人们有关的所谓的事实有时候被用作如下哲学理论的证据，这种理论主张自由意志和道德责任实际上是与因果决定论不相容的。有时候它又被用来提出如下看法，即如果有一种哲学理论主张不兼容主义的反面，那么这种哲学理论就有一种与生俱来的重要责任，要证明虽然我们的直觉站在兼容主义的对立面上，但是兼容主义仍然是正确的。

在一项关于自由意志和道德责任的极具影响力的研究中，艾迪·纳米亚斯(Eddy Nahmias)、斯蒂芬·莫里斯(Stephen Morris)、托马斯·纳德尔霍弗(Thomas Nadelhoffer)和简森·特纳(Jason Turner)(2004 年、2005 年、2006 年)质疑了人们是天然的不兼容主义者这一假设。考虑一下如下这个小场景：

---

① 可见斯特劳森(1986 年)、凯恩(Kane)(1999 年)、奥康纳(O' Connor)(2000 年)、佩雷布姆(Pereboom)(2001 年)、平克(Pink)(2004 年)。

超级计算机

> 设想下一世纪我们发现了所有的自然法则，我们造了一台超级计算机，它可以从这些自然法则和当前世界万物的状态中精确地演绎出未来世界上会发生什么事情。它可以看 30
> 到世界现在之所是，并带着百分之百的准确性来预测它将来之所是。假设那种超级计算机确实存在，它观察着宇宙在公元 2150 年 3 月 25 日的某个时刻的状态，而这时就是杰里米·哈尔(Jeremy Hall)出生前的 20 年。计算机从这一信息和自然法则中推导出如下信息：杰里米一定会在 2195 年 1 月 26 号下午 6 点抢劫富达银行。超级计算机的预测像往常一样精准；杰里米在 2195 年 1 月 26 号下午 6 点抢劫了富达银行。

如果人们是天然的不兼容主义者，那么我们应该预测大多数被要求去考虑超级计算机例子的人都会判断杰里米抢银行并非出自他自己的自由意志，无须为他的行为负道德责任。然而，令人惊奇的是，纳米亚斯等人发现事情刚好相反。他们发现，大多数被要求去考虑超级计算机例子的人都判断杰里米抢银行出自他自己的自由意志，要为他的行为负道德责任。① 这一点揭

---

① 更具体地说来，在那些被要求对**超级计算机**例子中杰里米的行为作出评价的实验对象中，有 72%的实验对象判断杰里米抢银行出自他的自由意志，而有 83%的实验对象判断他的行为在道德上是**应受指责的**。有趣的是，似乎实验对象并没有受到杰里米的行为的负面性质的影响。实验对象被要求去考虑一个相似的例子，在其中杰里米救了一个小孩的生命，其中 68%的实验对象判断杰里米的行为是出自他自己的自由意志，而有 88%的实验对象认为他应该为他的行为受道德上的**褒扬**。为了使因果决定论足够明确，纳米亚斯等人考察了另外两个场景：在一个场景中，一个宇宙被一次次地重新创造出来，带着相同的初始条件，在另一个场景中，人们的行为完全由基因和环境的因素所决定。在两个例子中，大多数被要求去评价这个行动者在这个例子中的行为的实验对象都判断这个行动者是出自他或她的自由意志而行动的，即使这些行为是被因果地决定的，同时这个行动者要为他或她的行为负道德责任。

示了人们实际上可能是**天然的兼容主义者**，他们相信自由意志和道德责任实际上是相容于因果决定论的。如果人们是天然的兼容主义者，那么这意味着兼容主义应该被给予一个缺省的、正面的重要认知地位，而不兼容主义则有一种与生俱来的责任，要证明虽然我们的直觉站在它的对立面上，但是不兼容主义仍然是真的。

尽管这意味着在关于自由意志和道德责任的哲学讨论中出现了一个命运的逆转，肖恩·尼科尔斯和约书亚·诺布（2007年）提出，我们可以把这些结论和人们是天然的不兼容主义者这一看法调和起来，其办法就是密切关注**情感**在产生相关哲学直觉中产生的作用。根据尼科尔斯和诺布的看法，即使人们是天然的不兼容主义者，当他们面对一些充满情感的小场景时，也可以形成兼容主义的直觉。这里的基本看法是，人们对充满情感的小场景的情感反应扭曲了他们关于道德责任到底是不是与因果决定论相容的判断，将他们从天然的不兼容主义带向了**不天然的兼容主义**。如果这是真的，那么似乎我们便可以解释如下这个事实，即大多数被要求去考虑**超级计算机**例子的人都判断杰里米的行为是出自他自己的自由意志的，是要负道德责任的，
31 同时无须抛弃人们是天然的不兼容主义者这一看法。

但是，这个看法对不对呢？人们是天然的不兼容主义者吗？他们的直觉判断被他们对充满情感的小场景的情感反应扭曲了吗？为了弄清楚事情到底是不是这样，尼科尔斯和诺布进行了另一项研究，这项研究既包含那种没有什么情感内容的抽象问题，也包括富含情感内容的具体例子。一开始，尼科尔斯和诺布为人们提供了对决定论和非决定论的如下描述：

> 设想一个世界（世界 A），在其中发生的一切完全由在

它之前发生的所有事情引起的。这个世界一开始就是这样的。因此,世界之初发生的事情导致了后来发生的事情,以此类推,直到现在。例如,有一天,约翰决定午餐吃炸薯条。像其他事情一样,这个决定完全是由之前发生的事情引起的。因此,如果在约翰作出他的决定之前,这个世界中的一切都完全不变,那么约翰决定吃炸薯条这件事将**不得不**发生。

现在设想另一个世界(世界 B),在这个世界中,**几乎**所有发生的事情都是完全由之前发生的事情引起的。而人类的作决定是个例外。例如,有一天玛丽决定午餐吃炸薯条。因为在这个世界里,一个人的决定并不是完全由之前发生的事情引起的。因此即便在玛丽作出她的决定之前这个世界中的一切都完全不变,玛丽决定吃炸薯条这回事**并不一定非得发生**。她可以决定吃点别的。

这里的主要区别是在世界 A 中每一个决定都是由这个决定之前发生的事情所完全地决定的——在过去被给定的情况下,每一个决定都**不得不**如它所发生的那样**发生**。相反,在世界 B 中,决定并非完全由过去引起,人类的每一个决定都**无须**如它所发生的那样**发生**。①

尼科尔斯和诺布发现,在读完这段对决定论和非决定论的

① 这里要简要地谈一谈这些研究是如何描述决定论的。纳米亚斯等人(2007年)认为人们的哲学直觉也受到决定论被描述的方式的影响。具体说来,比起用机械的、神经科学的术语来描述,当用非机械的、心理性的词汇来描述决定论时,人们更有可能表达出兼容主义的直觉。有趣的是,它们还发现人们对某个充满情感的小场景的情感反应影响了他们关于道德责任是否与因果决定论相容的直觉判断,不论这种决定论是如何被描述的。罗斯基(Roskies)和尼科尔斯(2008 年)也发现了人们的哲学直觉受到决定论被描述的方式的影响。他们发现,人们的哲学直觉敏感于这种决定论场景是作为现实的还是纯粹可能的例子被给出的。

32 描述后，大多数被问及“世界 A 中的某个行动者为她的行为负道德责任是否可能”这一问题的人都判断这是不可能的。① 换言之，大多数人在面对一个关于因果决定论和道德责任的兼容性的抽象问题时，都判断这两者实际上是不兼容的。但是，考虑一下下面的小场景：

蓄意谋杀的丈夫

在世界 A 中，一个名叫比尔的男人被他的秘书迷住了，而且他很确定与她在一起的惟一方法就是杀死他的妻子和 3 个孩子。他知道，在着火时要逃离他的房子是不可能的。在一次公事出差前，他在自家的地下室里做了点手脚，烧毁了这个房子，杀死了他的家人。

有趣的是，尼科尔斯和诺布发现，大多数被要求考虑**蓄意谋杀的丈夫**这一案例的人都判断比尔是要为杀害他的妻子孩子负道德责任的，即使他的行为是被因果地决定的。② 当面对抽象的、缺少情感的内容时，似乎人们是天然的不兼容主义者，但是当面对具体的、富含情感内容的例子时，他们就给出了兼容主义的直觉判断。③ 尼科尔斯和诺布承认，人们的情感反应对他们

---

① 更具体地说来，当被问及世界 A 中的某个行动者是否可能为她的行为负完全的道德责任时，有 86%的实验对象判断这是不可能的。

② 更具体地说来，当被要求去考虑**蓄意谋杀的丈夫**的例子时，有 72%的实验对象判断比尔要为他的行为负道德责任。

③ 为了消除这样一种可能性，即人们的判断仅仅取决于抽象或具体的区分，尼科尔斯和诺布又执行了第二个研究，这个研究只涉及高情感度(包含习惯性强奸)和低情感度(包含习惯性逃税)的具体例子。他们的结论证实了一点，即使某个人的行为是被因果地决定的，比起低情感度的例子，在高情感度的例子中人们更有可能说这个人要为她的行为负道德责任。

判断的扭曲解释了这个事实。这并不足以揭示不兼容主义是对的，但是它确实揭示了也许纳米亚斯等人过于心急地得出了人们是**天然的**兼容主义者这一结论。相反，我们可以说人们是**不天然的兼容主义者**。

然而，试图将观察到的这种偏差作为一种**执行错误**（performance error）而将它解释掉，[1]也是有问题的。[2] 要描述一个过程干涉了另一个过程，这预设着对所涉及的这些过程的理解。如果我们已经很好地说明了那些掌管着我们的**道德责任**判断的特定机制，或者已经刻画了那些掌管着这些判断的认知机制应当运算的函数，[3]那么，或许我们能够确定我们的情感反应是否干涉了那些机制的恰当作用——换言之就是我们的情感反应所带来的影响是否构成了一种执行错误。问题在于我们没有这样的东西。我们具有的是外延的分布数据：在 y 条件下， 33
在 x 场景中，百分之多少的实验对象给出了答案 z。但是这仅仅提供了一种对这种机制所运算的函数的输入/输出型的说明。[4] 它没有让我们了解这个运算过程是如何实施的，特别是没有让我们了解相关的表征和算法，也没有让我们了解这种运算过程在物理上是如何实现的。[5] 没有类似这种信息的东西，

---

① 这里的英文原文是“explain away”，翻译成“解释掉”也是勉为其难的，因为它的意思是将某个异常的东西放在一个理论框架之内，用一套理论来解释这个现象，这个现象就显得不再异常了。——译者

② 对这一点的进一步讨论，可见亚历山大等人（2010 年 a，2010 年 b）。

③ 根据心灵的运算理论，人类的心灵或大脑是一个信息处理系统，大脑中的各个“机制”是一个个能够独立运算的“模块”。此处“函数”的原文为“function”，翻译成“函数”，是咨询过作者约书亚·亚历山大的。——译者

④ 马尔（Marr）（1982 年）称此为**计算**层次（computational level），派利夏恩（Pylyshyn）（1984 年）称此为**语义学**层次，格拉斯（Glass）等人（1979 年）称此为**内容**层次。

⑤ 马尔（1982 年）将这两个层次称为**算法**层次和**运用**层次，派利夏恩将它们称为**句法**和**物理学**层次，格拉斯等人将它们称为**形式**和**方法**层次。

至少在这样的背景下，我们很难把导致能力的因子和导致执行错误的因子区分开来，尤其是当我们在处理一些大型的系统性变化时。因此，似乎尼科尔斯和诺布着实无法得出如下结论，即人们是天然的不兼容主义者，只不过被自身对富含情感的内容的情感反应带向了某种不天然的兼容主义。在没有对掌管着我们的**道德责任**判断的特定机制进行恰当地说明之前——他们似乎没有这样的说明，我们确实不能这样说。

到目前为止，我们还忽略了另一种可能性。也许**一些人**是天然的兼容主义者，而**另一些人**则是天然的不兼容主义者。比如，也许存在着多种**道德责任**的概念，依据其中的几种，自由和道德责任与因果决定论是兼容的，同时，依据另外的几种，自由和道德责任和因果决定论是不兼容的。这一可能性得到了如下事实的启示，即经验性研究证明了一些假定的例子引起了几种**不同的**直觉反应。同时，最近的一些经验性研究揭示了个性的差别可以对人们到底是天然的兼容主义者还是天然的不兼容主义者这一点产生影响（可见菲尔兹［Feltz］和科克雷［Cokely］，2009 年），这些经验性研究也支持了前述那种可能性。这一可能性还有一个优点，那就是它绕开了尼科尔斯和诺布的不天然的兼容主义观点所引发的难题。由于我们不再被要求用执行错误来解释掉那些被观察到的偏差，那么我们无法说清楚我们对富含情感的内容的情感反应是不是我们作出**道德责任**判断的能力的一部分这一点就变得不那么重要了。

为了弄清楚是不是一些人是天然的兼容主义者而另一些是天然的不兼容主义者，菲尔兹、科克雷与纳德尔霍弗（2009 年）等人开展了另一项研究，这项研究要求人们去考虑富含情感内容的小场景**和**缺乏情感内容的小场景。具体说来，他们要求人
34

们去考虑如下两个从尼科尔斯和诺布那里(2007 年)借用来的小场景：

强奸惯犯

[在一个决定论的世界中，]就像过去常常做的那样，比尔尾随了一个陌生人，并强奸了她。

习惯性逃税

[在一个决定论的世界中，]就像过去常常做的那样，马克又逃税了。

有趣的是，菲尔兹等人发现大多数被要求考虑这两个例子的人要么判断比尔和马克都无须为其行为负道德责任，要么判断比尔和马克都要为其行为负道德责任。① 这些结论和他们称之为**多种概念观点**的看法相一致——这种观点认为一些人是天然的兼容主义者而另一些人是天然的不兼容主义者。此外，这些结论与我们已经考察过的其他观点并不一致。天然的不兼容主义将会预测大多数人将会判断比尔和马克都无须为其行为负道德责任，天然的兼容主义者将会预测大多数人将会判断比尔和马克都要为其行为负道德责任，不天然的兼容主义将会预测大多数人将会判断比尔要为其行为负道德责任(因为这个例子富含情感)而马克无须为其行为负道德责任(因为这个例子并不包含情感)。由于这些预测都没有实现，因此似乎我们有了更多

---

① 更具体地说来，在那些被要求去考虑这两个例子的实验对象中，有 67%的人判断比尔和马克都无须为其行为负道德责任，25%的实验对象判断比尔和马克都要为其行为负道德责任，8%的实验对象给出了一人要负责另一人无须负责的答案。

的理由来怀疑这些观点。

多种概念观点自身并不是没有问题的。具体说来,它引发了一个棘手的困境。[①] 哲学家感兴趣的是从关于人们共同的直觉判断的主张过渡到关于哲学理论的认知地位的主张。但是,如果一些人是天然的兼容主义者,而另一些人是天然的不兼容主义者,那么我们就不应期望能发现共同的直觉判断了。剩下有两种选择。一方面,哲学家可以在那些由现实或假定的特定例子所产生的不同直觉判断中选择一种。然而,哲学家认为这
35 种选择有一个问题,那就是要解释为什么其他的直觉判断要被忽略。而且,正如分歧知识论(epistemology of disagreement)中越来越多的文献所证明的那样,在面对互相冲突的证据时,要确定该做点什么,这并非易事。[②] 另一方面,哲学家可以选择将他们的结论相对化。[③] 但是,虽然这种做法在某些例子中可能是恰当的,然而总的看来这一选择不可能被证明是有吸引力的。尤其是,从认为特定人群有着不同的思考事物的方式的**描述性**相对主义,过渡到认为思考事物的正确方式相对于特定人群的**规范性**相对主义,在历史上是有困难的。由于一种选择并不是那么简单,而另一种并不是那么有吸引力,因此多种概念观点为如下看法带来了一个重大难题,即我们可以简便而不成问题地从有关人们关于道德责任的哲学直觉的主张过渡到关于某些道德责任的哲学理论的真或表面合理性的主张。这个难题仍需得

---

① 对这一困境的其他讨论,可见亚历山大等人(2010 年 a)。

② 可见克里斯汀森(Christensen)(2007 年)、埃尔加(Elga)(2006 年)、费德曼(Feldman)(2006 年)、费德曼和沃菲尔德(Warfield)(2007 年)、凯利(Kelly)(2005 年,2007 年,2008 年)和怀特(White)(2005 年)。

③ 最近对相对主义的讨论,可见博加桑(Boghossian)(2006 年)和海勒斯(2006 年)。

到解答。[1]

这里,人们想要知道事情有了何种进展,也是很自然的。[2]起初,实验哲学家去关注关于自由意志和道德责任的性质的讨论,是因为这些讨论典型地包含了对哲学直觉的诉诸。兼容主义者和不兼容主义者都利用了这样一些论证,这些论证都是从关于人们的直觉判断的主张过渡到这样一些主张,这些主张是与关于自由、道德责任和因果决定论的兼容性的哲学理论的真或合理性相关的。实验哲学家希望能够提供一幅关于相关直觉自身的清晰画面,并提供对于这些心理机制以及塑造这些机制的那些影响的洞见,以此将这些讨论推向前进。一幅关于相关直觉的清晰画面将会帮助我们确定该接受哪些论证以及否定哪些论证。但是,现实比预料的更为复杂。我们不是有了一幅关于相关直觉的清晰画面,而是有了好几幅互相竞争的画面:也许人们是天然的兼容主义者,他们相信自由意志、道德责任与因果决定论是兼容的;也许人们是天然的不兼容主义者,他们被自身对于富含情感的内容的情感反应带向了某种不天然形式的兼容主义;也许不存在一组关于自由意志、道德责任和因果 36
决定论的单一直觉。由于我们还不知道哪一幅画面是正确的,因此我们无法确定该接受哪些论证以及否定哪些论证。此时此刻,我们确实无法将这些讨论推向前进。不过,重要的

---

① 尽管纳米亚斯和默里(Murray)(2011年)最近的工作给出了一些有趣的证据,这些证据可能提供了一些依据,让人们忽略一组关于道德责任的哲学直觉。他们发现大多数认为自由、道德责任与因果决定论不兼容的人之所以要这样做,是因为他们错误地相信了因果决定论包含某种**宿命论**,在这个相关的因果过程中,主体的精神状态(包括意图)被忽略了。

② 关于如何处理前面讨论的那些担忧,索默斯(2011年)给出了几种建议,并且指出了一些问题,随着关于这一主题的实验研究向前推进,这些问题是值得探讨的。

是要明白,对这一事实(即目前关于自由意志和道德责任的性质的哲学讨论无法以我们关于真实或假定的例子的直觉判断为依据而向前迈进)的认识本身就是一个进步,就是对当前关于自由意志和道德责任的哲学讨论的现状的一个重要贡献。毕竟,知道你至少目前无法使用某种工具,有时候这和找到正确的工具一样重要。

3. 实验哲学和知识论

当代知识论中有一个被争论了很久的问题,那就是知识的归属是否敏感于会话语境。一个悬而未决的难题就是如何理解如下两个小场景(德娄斯[DeRose],1992 年):①

银行的例子 A

在一个星期五的下午,妻子和我开车回家。我们计划回家路上去一趟银行,把我们的支票给存了。但是当我们开车路过银行的时候,我们发现里面的队伍特别长,在周五下午,队伍总是很长的。尽管我们通常都是希望尽快把支票给存了,不过在这个情况下,要立即把支票存了这一点并不十分重要。因此我建议我们直接开车回家,周六早上再来存支票。我妻子说:"也许明天银行不开门。很多银行周六都是关门的。"我回答:"不,我知道它会开门。两星期前的那个周六我就在这里。它一直开到了中午。"

---

① 在这场争论中,人们经常诉诸的其他小场景,可见科恩(Cohen)(1998 年)、方特(Fantl)和麦克格雷斯(McGrath)(2002 年)以及斯坦利(2005 年)。

银行的例子 B

> 在一个星期五的下午，妻子和我开车经过一家银行，就
> 像在银行的例子 A 中那样，我们注意到队伍很长。我还是 37
> 建议我们周六再存支票，并解释说两个星期前的那个周六
> 我就在这里，发现它一直开到中午。但是在这个例子中，我
> 们刚刚签收了一张面额很大、很重要的支票。如果我们的
> 支票在下周一早上之前不存进我们的账户，这张重要的支
> 票将会跳票，让我们处于一个非常糟糕的处境。当然，银行
> 在周日是不会开门的。妻子提醒我注意这些事实。她说：
> “银行会改变它们的营业时间。你知道这家银行明天会开
> 门吗？”尽管我和以前一样确信这家银行明天会开门，但是
> 我还是回答：“好吧，我不知道。我最好进去问问清楚。”

争论的一方是**语境主义者**，他们提出，我们对于这些以及与此类似的小场景的直觉反应揭示了某个会话语境的不同特征——具体说来就是哪种可能性被弄成是显著的（salient）——是可以影响知识归属的真值的。[①] 争论的另一方，则是**主体-敏感的不变主义者**（subject-sensitive invariantists），他们提出，会话语境不会影响知识归属的真值，我们对这些以及与此类似的小场景的直觉反应最好用其他方式来解释，比如可以用此人（被描述的正是此人的认知状态）所面临的风险来解释。[②]

让实验哲学家对这场争论尤其感兴趣的是，争论双方通常都诉诸关于同一些假定例子的同一些直觉。这里尚有争论的并

---

① 可见德娄斯（1992 年）、刘易斯（1996 年）和科恩（1998 年）。

② 比如，可见方特和麦克格雷斯（2002 年）、霍桑（Hawthorne）（2004 年）、斯坦利（2005 年）。斯坦利将他的不变主义称为**利益相关的不变主义**（interest-relevant invariantism）。

不是相关的直觉是什么,而是如何最好地解释它们:它们是为显著性的重要性还是为风险的重要性提供了证据呢?① 但是,这并不是实验哲学家近来对两者的争论感兴趣的惟一理由。正如乔纳森·谢弗(Jonathan Schaffer)(2005 年)所指出的,使这一争论极其有趣的东西,同时也使它极其难以解决:这一争论通常采用的那些假定例子在何种可能性被搞成是显著的*以及*在风险方面都会有所不同。这意味着相关的直觉确实无法告诉我们什么是我们想知道的东西,换言之,它们无法告诉我们哪一个(显著性还是风险)是重要的。这为实验哲学家对两者的争论感兴趣提供了另一个理由。实验哲学家也许可以仔细研究那种合
38 适的小场景,即那种把风险是什么与何种可能性被搞成是显著的这两者区分开来的小场景,并以此来将这些讨论推向前进。

让我们先来考察一下认为显著性是重要的这一看法吧。根据很多语境主义者的看法,某人是否知道某些东西(或者,至少被认为知道某些东西)这一点取决于,至少部分取决于何种可能性被搞成是显著的。语境主义者将他们的观点建立在我们关于前述两个银行的例子或其他有着类似结构的例子的直觉之上。在这两个例子中,丈夫和妻子在周五下午下班回家时途经一家银行,想把他们的支票给存了。当他们发现里面的队伍很长时,丈夫建议他明天回来存支票,并且对他的妻子说他知道银行周六会开门,因为他两个星期前的那个周六就在这里。这两个例

① 实际上,语境主义者和主体-敏感不变主义者的争论还要更复杂一些,因为这两群人都赞同知识敏感于主体面临的风险的影响。他们不一致的地方在于,把主体弄错的可能性搞成是显著的,这一点是否会影响认知标准。语境主义者认为它会影响到认知标准;而不变主义者认为它不会。这样说来,这场争论并不是真的与显著性重要还是风险重要这个问题有关(因为两者都赞同风险确实是重要的),而是与显著性到底是否重要这一问题有关。

子的惟一重要的区别就是，在第二个例子中，和第一个例子不一样的是，妻子明确地提出了丈夫可能会弄错的可能性，并指出银行有时候会改变它们的营业时间。语境主义者认为我们的直觉将会服从这一区别，具体说来，比起**银行的例子** B，我们在**银行的例子** A 中将会更乐意于说这个丈夫知道银行周六会开门。我们直觉的这种变化应该被认为是这样一种证据，它证明显著性是重要的，同时还证明语境主义是正确的。

这一论证是以一个可被检验的经验性主张为前提的，即我们的直觉服从于哪种可能性是显著的，为了测试这一主张，近来研究者们开展了一些实验研究，利用了一些小场景，确保各个小场景之间的惟一区别就是哪种可能性是显著的。令人惊奇的是，一些研究的结论揭示了显著性**并不**重要：与出错的可能性未被提及的情况下一样，在出错的可能性相当显著的情况下，人们似乎也愿意说某人是知道某物的。

比如，考虑一下下面这两个小场景（布克瓦尔特[Buckwalter]，2010 年）：

低标准

一个周五的下午，席尔薇和布鲁诺下班后开车回家。他们打算去银行把支票给存了，但是当他们开车路过银行 39
的时候，他们发现里面的队伍很长。尽管他们通常想要尽早把支票存了，然而这个例子中，他们是不是马上将它存掉并没有太大的关系。布鲁诺对席尔薇说："上个星期我就在这里，我知道银行周六也会开的。"布鲁诺建议他们先回家，周六再来存支票。当他们周六早上回来时，银行确实开门营业。

> **高标准**
>
> 一个周五的下午，席尔薇和布鲁诺下班后开车回家。他们打算去银行把支票给存了，但是当他们开车路过银行的时候，他们发现里面的队伍很长。尽管他们通常想要尽早把支票存了，然而这个例子中，他们是不是马上将它存掉并没有太大的关系。布鲁诺对席尔薇说："上个星期我就在这里，我知道银行周六也会开的。"布鲁诺建议他们先回家，周六再来存支票。席尔薇说："银行一般周六都是关门的。也许这家银行明天也不会开门。银行经常改变它们的营业时间；我记得这家银行过去的营业时间不是这样的。"当他们周六早上回来时，银行确实开门营业。

与**银行的例子A**和**银行的例子B**不一样的是，这两个例子间的惟一区别在于在**高标准**例子中布鲁诺搞错的可能性被明确地提了出来。如果语境主义是对的，那么比起那些被要求去考虑**高标准**例子的人，我们应该预期那些被要求考虑**低标准**例子的人更有可能判断布鲁诺知道银行周六会开门。但情况并不是这样。布克瓦尔特(2010年)发现，关于这两个例子的直觉判断之间没有什么重要的区别：和被要求去考虑**低标准**例子时一样，当人们被要求去考虑**高标准**例子时，他们给出布鲁诺知道银行周六开门这一判断的可能性是一样的。①

也许只有风险重要的时候，显著性才是重要的。布克瓦尔
40 特的这两个小场景所描述的都是这样的例子，在这些例子中，并

① 更具体地说来，实验对象被要求对赞同或反对布鲁诺知道银行周六开门的程度进行评分(使用5分制李克特量表法，1分＝强烈反对，5分＝强烈赞同)。那些被要求考虑**低标准**例子的实验对象给出的平均分数是3.83分，那些被要求考虑**高标准**例子的实验对象的平均分数是3.64分。

没有什么个人损失是与弄错联系在一起的;在这两个例子中,是不是立刻把支票存掉这一点并没有太大的关系。如果支票必须马上存掉,情况也许就会不同了。比如,考虑一下如下两个小场景(梅[May]等人,2010 年):

风险大且没有出错可能性

一个周五的下午,汉娜和她的妻子萨拉[1]下班后开车回家。她们打算回家路上去银行把支票给存了。由于她们的账单即将到期,而户头里又没有什么钱,因此她们在周六把支票存了这一点是**很重要的**。但是当她们开车路过银行的时候,她们发现里面的队伍很长,周五下午的队伍总是很长。汉娜说两星期前的那个周六的早上她就在这里,银行是开着的。汉娜说:"我知道银行明天会开门。我们可以明天再来存支票。"

风险大且有出错的可能性

一个周五的下午,汉娜和她的妻子萨拉下班后开车回家。她们打算回家路上去银行把支票给存了。由于她们的账单即将到期,而户头里又没有什么钱,因此她们在周六把支票存了这一点是**很重要的**。但是当她们开车路过银行的时候,她们发现里面的队伍很长,周五下午的队伍总是很长。汉娜说两星期前的那个周六的早上她就在这里,银行是开着的。**萨拉指出银行会改变营业时间的**。汉娜说:"我知道银行明天会开门。我们可以明天再来存支票。"

---

[1] 为了反对性别歧视,外国人写论文总喜欢用"她",而不是"他",大概由于这个原因,这里便让两个"她"结了婚。——译者

和布克瓦尔特那两个银行例子不同的是，这两个例子中的风险都是很大的。如果显著性比风险更重要的话，那么我们应该预测，比起那些被要求去考虑**风险大且有出错的可能性**的例子的人，那些被要求去考虑**风险大且没有出错可能性**的例子的人将更有可能判断汉娜知道银行明天会开门。但是，这一预测又没有实现。约书亚·梅(Joshua May)、沃尔特·辛诺特-阿姆斯特朗(Walter Sinnott-Armstrong)、杰·胡尔(Jay Hull)和阿伦·齐默尔曼(Aaron Zimmerman)(2010 年)发现，在关于这两个例子的直觉判断之间不存在什么有统计意义的区别。①

41 这些研究表明语境主义是错的。在上述研究中，研究者努力确保这几对小场景间的惟一区别就在于出错可能性是否被搞成是显著的。语境主义者预测人们的直觉判断将会追随这一区别，然而，和出错的可能性未被提及的情况一样，在出错的可能性被搞成显著的情况下，人们似乎也同样愿意说某人知道某物。似乎显著性**并不**重要。②

---

① 更具体地说来，实验对象被要求对赞同或反对汉娜知道银行周六开门的程度进行评分(使用 7 分制李克特量表法，1 分＝强烈反对，7 分＝强烈赞同)。那些被要求去考虑**风险大且没有出错可能性**的例子的实验对象给出的平均分是 5.07 分，而被要求去考虑**风险大且有出错的可能性**的例子的实验对象给出的平均分是 4.60 分。

② 对于这些研究，德娄斯(2011 年)提出了几个问题，包括在某些研究中，参与者被要求为他们**作出**知识归属的意愿进行评分，而不是被要求对知识归属进行**评价**，同时，当参与者被要求去评价知识归属时，他们甚至在如下这种情况下也被要求这样去做，在这些情况下，否定知识将会是更为自然的。由于什么样的会话语境是有关的这一点(至少部分地)取决于你是被要求去作出知识归属还是去评价知识归属，同时，由于在否定知识更为自然的情况下被要求去评价知识归属这一做法带来了一些实用主义的暗示(implications)，它们能对我们的判断产生影响，德娄斯担心这些研究实际上并没有在检验认知语境主义的为真性或表面合理性。这些问题是**合理的**，不过布克瓦尔特发现(即将出版)，即使在这些研究的实验设计中考虑到这些问题，显著性似乎也并不重要。关于实验哲学家所采用的方法，德娄斯还提出了一些更为一般化的问题；我们将在本章的结尾考察这些问题。

如果显著性并不重要,那么风险又怎么样呢?依据主体-敏感的不变主义者,某人是否知道某物依赖于(至少部分地依赖于)关于个人损失的事实,这种损失是与出错联系在一起的。这揭示了它与对知识的一种更为传统的考察(根据后者,与某人是否知道某物有关的惟一事实是有助于产生真[truth-conducive]的事实)的一个重要区别,并且是近年来知识论当中一种日益增长的反理智主义倾向的一部分。① 和语境主义者一样,主体-敏感的不变主义者将他们的看法建立在我们关于那几对小场景的直觉之上。比如,考虑一下如下两个小场景(斯坦利[Stanley],2005 年):

风险不大

一个周五的下午,汉娜和她的妻子萨拉下班后开车回家。她们打算回家路上去银行把支票给存了。因为她们没有什么要到期的账单,所以她们是不是这样做并不重要。但是当她们开车路过银行的时候,她们发现里面的队伍很长,周五下午的队伍总是很长。意识到她们是否立即把钱存掉这一点并不很重要,汉娜说:"我知道银行明天会开门,因为两星期前的那个周六的早上我就在这里。我们可以明天再来存支票。"

风险大

一个周五的下午,汉娜和她的妻子萨拉下班后开车回家。她们打算回家路上去银行把支票给存了。由于她们的账单即将到期,而户头里又没有什么钱,因此她们在周六把

① 还可以见方特和麦克格雷斯(2002 年),以及霍桑(2004 年)。

支票存了这一点是很重要的。汉娜说两星期前的那个周六
42 的早上她就在这里，银行是开门的。但是，萨拉指出，银行会改变它们的营业时间。汉娜说："我猜你是对的。我不知道银行明天会开门。"

假设银行周六会开门，假设在这两个小场景中，汉娜能够得到的事实（这些事实是有助于产生真的）保持不变，那么这两个例子之间惟一相关的区别应该就是与搞错联系在一起的个人损失了。因为这两个例子惟一相关的区别是与搞错相联系的个人损失，又因为主体-敏感的不变主义者认为我们对这些小场景的直觉将会是：在**风险不大**的例子中，汉娜知道银行周六会开门，而在**风险大**的例子中则不是这样。因此，他们得出结论，认为风险是重要的。

我们将会直觉地判断，在**风险不大**的例子中，汉娜知道银行周六开门，但是在**风险大**的例子中，汉娜则不知道。上述主张也是一个可以被检验的经验性主张。近来，实验哲学家做了一系列研究，其目的就是对这一主张进行检验，他们确保几对小场景之间的惟一差别就是与搞错联系在一起的个人损失。有趣的是，在这些研究中，有一些研究的结果揭示了风险**并不**重要：若各个小场景中的有助于产生真的事实是确定的，在搞错的个人损失较小**以及**较大的时候，人们都倾向于做出知识归属。

比如，考虑一下如下这对小场景（菲尔兹和扎彭提尼[Zarpentine]2010年）：

简化了的风险不大的例子

一个星期五下午，汉娜和她妹妹萨拉开车回家。她们打算回家路上去银行把支票给存了。因为她们没有什么要

到期的账单，所以她们是不是这样做并不重要。汉娜说两个星期前那个周六的早上她就在银行，银行是开着的。汉娜对萨拉说："我知道银行明天会开门。"

简化了的风险大的例子

一个星期五下午，汉娜和她妹妹萨拉开车回家。她们打算回家路上去银行把支票给存了。由于她们的账单即将到期，因此她们在周六把支票存了这一点是很重要的。汉娜说两个星期前那个周六的早上她就在银行，银行是开着的。汉娜对萨拉说："我知道银行明天会开门。" 43

这两个例子间的惟一区别是搞错的个人损失。如果风险是重要的，那么我们应该预料，比起那些被要求去考虑**简化了的风险大**的例子的人，那些被要求去考虑**简化了的风险不大**的例子的人更有可能判断汉娜知道银行周六会开门。有趣的是，事情好像并不是这样的。亚当·菲尔兹(Adam Feltz)和克里斯·扎彭提尼(Chris Zarpentine)(2010 年)发现，在关于这两个例子的直觉之间，不存在什么统计学上有显著性的区别。①

在另一个经验研究中，马克·费伦(即将出版)也发现了似

---

① 更具体地说来，实验对象被要求对赞同或反对汉娜知道银行周六开门的程度进行评分(使用 7 分制李克特量表法，1 分＝强烈反对，7 分＝强烈赞同)。那些被要求去考虑**简化了的风险大**的例子的实验对象给出的平均分是 3.85 分，而被要求去考虑**简化了的风险不大**的例子的实验对象给出的平均分是 3.83 分。菲尔兹和扎彭提尼(2010 年)还发现在人们关于类似的风险大或不大的例子的直觉判断之间不存在什么统计学上具有显著性的区别。事实上，当他们把他们所做的风险大或不大的实验统统放到一起时，他们确实发现了一个在统计学上具有显著性的区别，但是这影响的范围实在太小了。对于那些认为风险是知识归属的一个重要且无处不在的影响因素的哲学家来说，这应该不能带来什么安慰。

乎风险并不重要。考虑一下如下两个小场景：

不重要的

凯特在街上漫步，仅仅是散步，没有什么理由，也没有什么特定的地方要去。她走到了一个十字路口，并向一个路人询问这条街的名字。路人说："中央大道。"凯特看了看她的表，当时是早上11点45分。凯特的眼力相当正常，而且她清楚地看了她的手表。凯特的听力也相当正常，她很清楚地听到了路人的话。她没有什么特别的理由不相信路人的准确性。她也没有什么特别的理由不相信她的表的准确性。凯特可以去搜集关于她正站在中央大道上的进一步的证据（比如她可以去找一张地图），但是她没有这样做，因为，基于路人告诉她的东西，她已经认为她在中央大道上了。

重要的

凯特需要在中午的时候赶到中央大道：这关系到她的生计。她走到了一个十字路口，并向一个路人询问这条街的名字。路人说："中央大道。"凯特看了看她的表，当时是早上11点45分。凯特的眼力相当正常，而且她清楚地看了她的手表。凯特的听力也相当正常，她很清楚地听到了路人的话。她没有什么特别的理由不相信路人的准确性。
44 她也没有什么特别的理由不相信她的表的准确性。凯特可以去搜集关于她正站在中央大道上的进一步的证据（比如她可以去找一张地图），但是她没有这样做，因为，基于路人告诉她的东西，她已经认为她在中央大道上了。

和前面的例子一样，这两个例子之间的惟一区别是出错的个人损失。因此，如果风险是重要的，那么我们应该预期，比起那些被要求去考虑**重要的**这一例子的人，那些被要求去考虑**不重要的**这一例子的人将更有可能判断凯特知道她在中央大道上。但是，再一次地，情况似乎不是这样的。费伦在关于这两个例子的直觉判断之间并没有发现什么在统计学上具有显著性的区别。①

看来风险毕竟不是重要的。当我们把这些研究和关于显著性的实验工作放在一起的时候，这种结果挑战了语境主义和主体-敏感的不变主义。② 当出错的可能性显著的时候，和这种可能性不被提及的时候一样，人们似乎也一样愿意说某人知道某物，正如当与出错相联系的个人损失很大时，和这种损失较小时一样，人们似乎也一样愿意说某人知道某物。鉴于此，似乎应该否定语境主义和主体-敏感的不变主义，而应该支持某种形式的古典不变主义(classical invariantism)，按照这后一种看法，一旦人、命题和有助于产生真的事实被设定好了，那么知识归属的真值条件就不会改变。然而，我们有越来越多的理由来担心这一步可能走得太快了。比如，乔纳森·谢弗和约书亚·诺布(2011年)最近指出，那些关注显著性在知识归属中扮演的角色的实验研究并未成功地将银行周六会关门的可能性搞得足够显著。同时，N·安吉拉·皮内罗斯(N. Ángel Pinillos)(即将出版)最近

---

① 更具体地说来，实验对象被要求对凯特是否自信她正在中央大道上这一点进行评分(使用7分制李克特量表法，1分＝不自信，7分＝很自信)。那些被要求去考虑**不重要的**例子的实验对象给出的平均分是5.11分，而被要求去考虑**重要的**例子的实验对象给出的平均分是4.93分。

② 不过，布朗(Brown)(即将出版)最近提出大众的知识归属不敏感于诸如风险和显著性等因素这一点确实破坏了语境主义，但是没有破坏主体-敏感的不变主义。

也提出那些关注风险在知识归属中扮演的角色的研究并未确保参与者捕捉到正确的细节，具体说来，就是为了被视为知道某物而需要多少的证据和与搞错相联系的个人损失之间的关系。让我们更为仔细地审查一下这些观点。

根据谢弗和诺布的看法，仅仅提到一种可能性并不会必然
45 地让这一可能性变得显著，尤其是当这种被提及的可能性看上去相当奇怪或不大可能的时候。相反，这种可能性必须以一种相当具体和生动的方式呈现出来。这就对过去的实验研究提出了一个问题，那就是，在这些例子中，银行周六会关门的可能性并没有被搞得足够显著，若这种可能性被搞得足够显著了，那么显著性将会是重要的。为了检验这一预言，谢弗和诺布让实验对象考虑如下两个小场景：

修改了的低标准

一个周五的下午，汉娜和萨拉下班后开车回家。她们打算回家路上去银行把支票给存了。但是当她们开车路过银行的时候，她们发现里面的队伍很长，周五下午的队伍总是很长。汉娜说："两星期前的那个周六的早上我就在银行，银行是开着的。因此这是一家周六开门的银行。我们现在可以走了，明天再来存支票。"萨拉回答："好吧，听上去不错。我们周六再来。"

明确的高标准

一个周五的下午，汉娜和萨拉下班后开车回家。她们打算回家路上去银行把支票给存了。但是当她们开车路过银行的时候，她们发现里面的队伍很长，周五下午的队伍总是很长。汉娜说："两星期前的那个周六的早上我就在银

> 行，银行是开着的。因此这是一家周六开门的银行。我们现在可以走了，明天再来存支票。”萨拉回答：“哦，银行有时候会改变它们的营业时间的。有一次，银行改变了营业时间，周六关了门，我的兄弟里奥就为此陷入了麻烦。你想象一下，明天开车过来，发现大门紧闭。这该多么令人沮丧啊！”

基于更早的那些关注显著性在知识归属中扮演的角色的研究，我们应该预料在这里不会有什么区别：比起那些被要求去考虑**修改了的低标准**例子的人，那些被要求去考虑**明确的高标准**例子的人应该会同样判断汉娜知道银行周六会开门。但是，这不是谢弗和诺布发现的东西。他们发现，当汉娜搞错银行周末的营业时间的可能性以这种更为生动的方式被提出的时候，比起那些被要求去考虑**明确的高标准**例子的人，那些被要求考 46
虑**修改了的低标准**例子的人实际上更加倾向于判断汉娜知道银行周六将会开门。①

从这一点中，谢弗和诺布得出结论，认为显著性毕竟还是重要的，并且继续用这些结论（外加其他一些实验研究）来支持对比主义（contrastivism）。这是语境主义的一个版本，它认为关于某人知道某物的断言实际上等同于某人知道某物**而不是其他东西**的断言。但是，虽然这些结果无疑会让任何一些这样的人

---

① 更具体地说来，实验对象被要求对赞同或反对汉娜知道银行周六开门的程度进行评分（使用7分制李克特量表法，1分＝强烈反对，7分＝强烈赞同）。那些被要求去考虑**修改了的低标准**例子的实验对象给出的平均分是5.54分，而被要求去考虑**明确的高标准**例子的实验对象给出的平均分是3.05分。这一差别在统计学上达到p值小于0.001的显著水平。内格尔（2010年）也发现显著性是重要的，他发现把出错的可能性搞得显著起来，会使人们更不可能赞同对知识的断定，而更有可能赞同对知识的否定。

感到踌躇，这些人可能会倾向于以布克瓦尔特（2010 年）和梅等人（2010 年）的研究结果为基础来反对语境主义，不过谢弗和诺布回到语境主义，至少回到语境主义的某个版本的做法似乎太快了点。首先，由于谢弗和诺布只检验了澳大利亚的实验对象，这也许只是跨文化多样性的证据：也许显著性**只是对澳大利亚人来说**才是重要的。（正如我们在第四章中将会看到的那样，这并不是一个毫无根据的担忧。基于直觉的跨文化多样性的发现是实验哲学最重要的贡献之一。）其次，比起**高标准**的例子，在**明确的高标准**例子中出错的可能性是不是更加显著，这一点实际上也是不清楚的。在这两个例子中，主人公均被告知，银行有时候会改变它们的营业时间，这一事实被用来强调这样一种可能性，即正在讨论的这个特定的银行周六可能不会开门。事实上，似乎这两个例子之间的惟一重要的差别就在于，在**明确的高标准**例子中，主人公被要求去设想一下与出错相联系的个人损失。如果这是对的，那么即使我们不去理会那些关于跨文化多样性的担忧，谢弗和诺布似乎也得出了错误的结论。如果这些例子为什么东西提供了证据，那么它们似乎为**风险**的重要性提供了证据。

风险重要这一看法也得到了 N. 安吉拉・皮内罗斯（即将出版）最近的工作的支持。比如，考虑一下如下两个小场景（皮内罗斯，即将出版）：

风险不大的拼写错误

彼得是一名优秀的大学生，他刚刚为英语课写完了一
47 篇两页的文章。文章明天就要上交了。即使彼得的拼写是很好的，他也随身带有一本字典，他可以用它来检查拼写，确保没有什么拼写错误。但是这并没有什么关系。老师只要一个粗略的大纲，有少数几个拼写错误并无大碍。即便

这样，彼得也希望没有任何的错误。

风险大的拼写错误

约翰是一名优秀的大学生，他刚刚为英语课写完了一篇两页的文章。文章明天就要上交了。即使约翰的拼写是很好的，他也随身带有一本字典，他可以用它来检查拼写，确保没有什么拼写错误。这是至关重要的。老师是一个较真的人，如果有一个拼写错误，就不可能得到A。他需要完美。然而，约翰发现他自己处在一个不同寻常的处境中。为了在班上得到A，他的这篇文章需要得到A。为了保住他的奖学金，他需要在班上得到A。没有奖学金，他就不能呆在学校里了。离开大学将会让约翰和他的家人精神崩溃，为了让约翰上学，他的家人已经牺牲了很多东西。因此，对约翰来说，文章中没有拼写错误是极其重要的。他很明白这一点。

皮内罗斯问一些人：彼得或约翰需要校对这篇文章多少次，才知道那里已经没有什么拼写错误了。如果风险不重要，那么我们应该预期人们认为约翰在知道文章中没有拼写错误之前要校对文章的次数和彼得要校对文章的次数之间没有什么大的差别。这不是我们所发现的；相反，人们认为，为了知道文章中已经没有拼写错误，约翰要校对文章的次数平均要比彼得多出三倍。① 将这些结论和谢弗及诺布的结论放在一起，似乎风险

① 更具体地说来，那些被要求考虑**风险不大的拼写错误**的例子的实验对象的平均反应是2次，那些被要求考虑**风险大的拼写错误**的例子的实验对象的平均反应是5次。这一差别在统计学上达到p值小于0.001的显著水平。

毕竟还是重要的。然而，即使这些结论也是富有争议的，韦斯利·布克瓦尔特(Wesley Buckwalter)(即将出版)提出，比起关于风险和知识归属，这些结论更多地揭示了一些关于风险和理性行为之间关系的东西。①

我们从所有这些研究结论中得到了什么呢？一方面，我们
48 有经验证据表明风险和显著性都不重要。当搞错的可能性显得显著的时候，和这种可能性不被提及的时候一样，人们似乎都愿意说某人知道某物；在与搞错相联系的个人损失很大的时候，和这种损失不大的时候一样，人们似乎也都愿意说某人知道某物。另一方面，我们也有经验证据表明**某种东西**是重要的——但是，现在我们无法决定到底风险还是显著性是重要的。把事情搞得愈加复杂的是，我们不知道风险是否对知识归属的真有一种认知上的影响，还是仅仅对某些种类的行为的合理性具有一种实际的影响。② 目前，要从任意一组经验证据中获得结论，我们还须谨慎。③ 也许更早的研究没有发现什么在统计学上具有显著性的差别，这是因为根本不存在什么差别；也许它们找不到什么在统计学上具有显著性的差别，这是因为它们用来测试这种差别的工具不够精密或没有得到正确的调试。也许后来的研究发现了在统计学上具有显著性的差别，因为比起那些在先前研究中被使用的工具，它们用来测试那种差别的工具更加精密或得

① 皮内罗斯和辛普森(Simpson)(即将出版)用一系列更直接关注知识归属的新研究对这一担忧做出了回应，这些研究仍然发现了有关风险的重要性的证据。

② 见霍桑和斯坦利(2008 年)对知识和理性行为之间的关系的有趣讨论。

③ 威瑟森(即将出版)提出那些证明风险重要的貌似非决定性的证据实际上已经是决定性证据。按照威瑟森的说法，主体-敏感的不变主义者并不认为风险总是重要的；相反，他们只是认为风险有时候是重要的。如果这是对的，那么我们应该不用过分担心如下事实，即一些实验研究揭示了风险是重要的，而另一些研究则揭示了它并不是重要的。

到了正确的调试;也许后面的研究发现了在统计学上具有显著性的差别,因为这些研究所考察的是另一个认知人群。也许后来的研究发现显著性是重要的;也许后来的研究发现风险是重要的。也许风险对于知识来说是重要的;也许风险只对理性行为来说才是重要的。在这一点上,我们真的不知道情况是怎么样的。这需要进一步的研究。同时,对于那些用我们关于银行例子的直觉判断来提出某些哲学论证和观点的哲学家来说,这一点本身就有着重要的意义。不管我们有什么其他理由来接受语境主义或主体-敏感的不变主义的某些版本,在我们的哲学直觉能够有助于解决这两者间的争论之前,似乎需要更多的实验工作。

4. 结　论

在关于知识和道德责任之本性的当代讨论中,我们的哲学
直觉扮演着一个重要的角色,实验哲学家希望那些经验工作能 49
够提供一幅关于相关哲学直觉的清晰画面,借此将这些讨论推向前进。结果发现,这幅画面并不是那么清晰。在我们希望发现一致性的地方,我们发现了不一致性。我们可以接受这一点,或者解决它,或者将其解释掉;但是不管我们选择哪一种做法,都需要更多的经验工作。对于实验哲学(它的核心思想是:对我们哲学直觉以及所涉及的认知过程的理解是一件需要持续而仔细的经验考察的事)来说,认识到这一点本身就是某种胜利。认识到这种考察尚未完成,并不是说我们不需要它,而只是明白了为什么我们需要更多。

第三章
实验哲学和心灵哲学

1. 引　　言

那些对哲学直觉能够告诉我们的*东西*感兴趣的哲学家欠我 50
们一个用来解释哲学直觉*如何*能够告诉我们一些关于我们自身和我们周遭世界的东西的陈述。正如阿尔文·戈德曼(2007年)提出的，对于那些对哲学直觉关于这个非心理的世界可能告诉我们的东西感兴趣的哲学家来说，这个陈述可能是很难给出的。① 问题并不在于哲学直觉是精神状态；知觉也是精神状态，而我们不必去担心它们告诉我们关于这个非心理的世界的东西的能力。相反，问题在于，关于那种机制，关于那种可以将我们的哲学直觉与这个非心理的世界(哲学直觉应当告诉我们一些关于这个世界的东西)联系起来的机制，我们缺乏一种可被自然地接受的模型。这种担忧使得一些哲学家将他们的注意力转向了哲学直觉关于我们自身能够告诉我们的东西。根据这一看法，也许哲学直觉不能告诉我们世界的真实样子，但是它们可以告诉我们一些重要的东西，这些东西与我们思考这个世界的方式有关——具体说来，就是我们个人的或共有的概念。

在一些更为传统的圈子中，这一兴趣，即对哲学直觉关于我 51

们个人的或共有的概念能够告诉我们的东西的兴趣，是与**概念分析**联系在一起的。概念分析旨在确定哲学概念的精确意义，通常包括用我们关于那些概念的使用条件的哲学直觉来衡量哲学概念的某些定义。比如，分析哲学家可能会利用我们的这样一种直觉——这种直觉是与恰当地以及不恰当地将知识归属给某人有关的——来衡量各种可能的**知识**定义，并试图确定"知识"这一概念的精确定义。

在实验哲学的圈子里，这种对哲学直觉关于我们自身可能告诉我们的东西的兴趣有着另外一副模样。很多实验哲学家对确定哲学概念的精确意义不那么感兴趣，他们更为感兴趣的是确定那些影响我们对这些概念的使用的因素。因此，比起确定"知识"这一概念的精确意义，实验哲学家更感兴趣的是确定为什么我们会有那些关于恰当地以及不恰当地将知识归属给某人

---

① 还可以见戈德曼和普斯特(1998年)。戈德曼的担忧类似于现在以"贝纳塞拉夫问题"* 而为人所知的那个例子：根据我们的最好的数学真理理论，数学陈述指向抽象的实体；但是，根据我们的最好的知识理论，一个人知道 p，当且仅当那个人和使得 p 为真的东西之间有着一种因果的联系；因为抽象实体无法处于一种因果关系中，这就有了一个问题。这个担忧使戈德曼接受了一种"**心灵主义**"，这种观点认为哲学直觉只是与我们的概念的某些事实(比如概念的结果或意义)有关的证据。

* 原文为"Benacerraf's Problem"，保罗·约瑟夫·所罗门·贝纳塞拉夫(Paul Joseph Salomon Benacerraf)是一位出生于1931年的美国哲学家。这里指的应该是贝纳塞拉夫的认识论问题。贝纳塞拉夫针对科学哲学中一系列柏拉图主义的立场构造了一个问题。这个问题的论证尤其针对的是诸如哥德尔等人对数学直觉问题的考察。贝纳塞拉夫的论证从这样一个前提开始，即我们的最佳知识理论是知识的因果理论。因此，根据柏拉图主义，抽象对象并不坐落在时空中，而活生生的数学家则坐落在时空中。我们的最佳认识论理论告诉我们关于数学实体的知识应该来自于这些实体之间的因果性的相互作用。但是我们很难想象这会是真的。(译自"斯坦福哲学大全"网站)贝纳塞拉夫在1973年的《哲学期刊》(*The Journal of Philosophy*)上发表了一篇名为《数学真理》的文章，节选其中原文：我赞同一种知识的因果理论，根据这种理论，为了让"X 知道 S"是真的，这需要在 X 和 S 的名称、谓词和量词所指示的东西之间确定某种因果关系……但是，将这个关于知识的观点和数学真理的"标准"看法结合起来，就使得我们很难看清数学知识是如何可能的……数论中一些陈述的真值条件和任何相关的事件(这些事件与那些应该具有数学知识的人有关)的关系没办法获得理解。(第671—673页)——译者

的直觉。我们的目标不是,至少不仅仅是用我们的直觉来解释我们的哲学概念的意义,而是去解释我们的直觉本身。

在本章中,我们将会考察这类实验哲学的一个例子,这个例子特别具有影响力,那就是最近的这样一些实验工作,这些工作是关于规范性考虑(normative considerations)对人们思考世界的日常方式带来的显著影响的。在这种工作中,有很大一部分工作的着眼点是:规范性考虑是如何施加这种影响的,以及它们是构成了我们的**概念能力**的一部分,还是仅仅在我们的**概念运用**中扮演着某种角色。结果发现,要确定规范性考虑是不是我们的概念能力的一部分,这是很困难的,其部分原因在于要精确地确定规范性考虑是如何影响我们思考世界的日常方式的,这被证明是十分困难的。但是,正如休谟(Hume)多年前告诉我们的那样,即使知道它们是如何影响我们思考世界的日常方式的,这也不足以告诉我们它们是否应该具有这样一种影响。同时,我们从这项工作中获得的一个重要教益,那就是在我们能够确定规范性考虑是我们的概念能力的一部分还是仅仅在我们的概念运用中扮演着某种角色之前,还需要更多的(以及不同的)工作。

2. 副作用效应

有时候,人们关于一个特定行为是故意或不是故意做出的 52
直觉受到了**规范性**考虑的影响。也就是说,当人们在评价一个给定的行为是不是故意做出的时候,他们的判断受到了他们关于这一行为(或这一行为的结果)在道德上是善还是恶的信念的影响。当我们考虑包含**预知的副作用**(foreseen side effects)的例子时,这一点就会变得尤为明显。(一个结果是预知的副作

用,仅当一个行动者执行了一项行为,她预见这项行为将会导致某个结果,而她并不是为了导致这一结果而执行这项行为的。)比如,考虑一下如下两个小场景:

危害环境

一位公司的副总裁到董事长那里,说:"我们正在考虑开始一项新的计划。它将帮助我们增加利润,但它也会危害环境。"董事长回答:"我才不管什么危害环境。我只想尽可能获得更多的利润。让我们开始这项计划吧。"他们开始了这项新的计划。环境果然被破坏了。

改善环境

一位公司的副总裁到董事长那里,说:"我们正在考虑开始一项新的计划。它将帮助我们增加利润,同时它也会改善环境。"董事长回答:"我才不管什么改善环境。我只想尽可能获得更多的利润。让我们开始这项计划吧。"他们开始了这项新的计划。环境果然得到了改善。

实验哲学中有一个最著名的研究结果,那就是约书亚·诺布(2003 年 a)发现大多数被要求考虑*危害环境*例子的人都判断这个总裁是故意地破坏环境的,而大多数被要求考虑*改善环境*
53 例子的人都判断这个总裁不是故意地改善环境的。① 情况似乎

① 更具体地说来,在被要求评价总裁在*危害环境*例子中的行为的实验对象中,有 82%的人判断总裁是故意破坏环境的,而被要求评价总裁在*改善环境*例子中的行为的实验对象中,只有 23%的人判断总裁是故意改善环境的。这一差别在统计学上达到 p 值小于 0.001 的显著水平。诺布(2003 年 a)发现,从另一对有着相同的基本结构的小场景中,也可以得出类似的结果。

是这样的，比起那些人们认为副作用在道德上为善的情况，当人们认为副作用在道德上为恶时，他们更倾向于判断一个行动者是故意导致某个副作用的。这为如下这一观点提供了证据，即人们关于某个行为（或那一行为的结果）的道德状况的信念影响到了他们关于这一行为是不是故意做出的直觉。

道德考虑对人们的故意性判断的显著影响并不仅仅局限于涉及副作用的例子。比如，相同的影响似乎也可以在涉及这样一种行动者的例子中出现，这一类行动者试图导致某一结果，然而却缺乏可靠地导致这一结果所必需的技巧。让我们再来考虑一下这对小场景：

没有技巧/正面结果

杰克很想赢得步枪比赛。他知道，只有击中靶心才能赢得比赛。他举起步枪，瞄准靶心，按下了扳机。但是，杰克并不擅长使用他的步枪。他手在枪膛上滑了一下，胡乱放了一枪……然而，这颗子弹直接落在了靶心上。杰克赢得了比赛。

没有技巧/不道德的结果

杰克很想得到更多的钱。他知道他阿姨死后他会继承一大笔钱。有一天，他看到阿姨经过窗前。他举起他的步枪，瞄准她，然后按下了扳机。但是，杰克并不擅长使用他的步枪。他的手在枪膛上滑了一下，胡乱放了一枪……然而，这颗子弹直接击中了她的心脏。她马上死了。

诺布(2003 年 b)发现大多数被要求考虑**没有技术/不道德结果**例子的人判断杰克是故意打中他阿姨的，而大多数被要求考虑**没有技巧/正面结果**例子的人则判断杰克并非故意打中靶心。① 比起那些道德上中立的结果，如果人们认为某个结果在道德上是恶的，那么人们就更倾向于判断一个行动者是故意导致这个结果的。这再一次为如下这个观点
54 提供了证据，即人们关于某个行为(或该行为的结果)的道德状态的信念影响了他们关于这个行动是不是故意做出的直觉。

前面考察的这些研究指明了人们的道德判断可以影响他们关于某个行为是不是故意做出的直觉。但是，这意味着什么呢？这儿有两个主流解释。根据第一个解释，这些结论告诉我们一些东西，这些东西是和道德判断与人们对“故意的”一词的使用的关系有关的，这些结论最好用那些掌管着我们的语言实践的潜在机制的具体特征来解释(比如，可见亚当斯[Adams]和斯特德曼[Steadman]，2004 年 a，2004 年 b)。根据第二个解释，这些结论告诉了我们一些非常具体的东西，这些东西是和道德判断与人们的“故意行为”概念的关系有关的，这些结论最好用掌管着人们关于故意行为的判断的潜在机制的具体特征来

---

① 更具体地说来，在那些被要求评价**没有技巧/不道德结果**例子中的行为的实验对象中，有 76%的人判断他是故意打中他阿姨的，而在那些被要求评价杰克在**没有技巧/正面结果**例子中的行为的实验对象中，有 28%的人判断他是故意打中靶心的。这一差别在统计学上达到 p 值小于 0.01 的显著水平。除去刚才讨论的那两项研究外，库什曼和米尔(2008 年)、诺布(2004 年 a，2004 年 b)、诺布和伯拉(Burra)(2006 年)、诺布和曼德卢(Mendlow)(2004 年)以及莱斯利(Leslie)等人(2006 年)、马勒(2006 年)、麦肯(McCann)(2005 年)、米尔和库什曼(2007 年)、纳德尔霍弗(2004 年 a，2004 年 b，2005 年，2006 年 a，2006 年 b)以及杨等人(2006 年)都发现在有着相同基本结构的不同的小场景中可以得出类似的结论。

解释(比如,可见诺布,2003 年 a,2006 年,纳德尔霍弗,2004 年 a,2004 年 b,2006 年 b,尼科尔斯和乌拉托夫斯基[Ulatowski],2007 年,以及马舍雷,2008 年)。[①] 让我们来依次考察这两个解释。

3. 副作用和会话语用学

弗雷德·亚当斯(Fred Adams)和安妮·斯特德曼(Annie Steadman)是最早试图解释副作用的那一批人。亚当斯和斯特德曼(2004 年 a)提出,我们应该对如下这一点持谨慎态度,即用关于人们的直觉判断的事实来获得关于人们的潜在概念(underlying concept)的性质的结论。具体说来,他们提出,关于我们的“故意行为”的大众概念,副作用并未告诉我们任何东西,相反,副作用最好完全用语用学的考虑来解释——具体说来就是用**会话的暗示**(conversational implicature)来解释。[②] 根据亚当斯和斯特德曼的说法,在**危害环境**的例子中,人们相信这个总裁是应受谴责的。这个信念引入了一个问题。人们用来表达这一信念的惟一办法就是主张总裁是故意破坏环境的。如果他们说总裁并不是故意破坏环境的,那么这将意味着这个总裁不必为破坏环境而受到谴责。认识到这一点后,当人们被问及这 55
个总裁是不是故意破坏环境时,他们倾向于说他是故意这样做

① 最近出现了第三个解释,即这些结论更一般地告诉了我们一些关于道德判断和**大众心理**的关系的东西,并且最好用这样一些潜在机制的具体特征来解释,这些机制掌管着多种大众心理判断。可见诺布(2006 年),纳德尔霍弗(2006 年 b),艾利克(2008 年),纳多(2008 年),以及佩蒂特和诺布(2009 年)。

② 格莱斯(Grice)(1989 年)引入了会话的暗示这一概念,用来指这样一些东西,这些东西是被一个句子所暗示出来的,而且这些东西不同于这一句子的字面意思所表达或断定的东西。

的。这种说总裁故意破坏环境的倾向并未反映人们的“**故意行为**”的大众概念的本性,它仅仅反映了他们想要因环境被破坏而谴责这个总裁的愿望。①

诺布(2004 年 b)提出,衡量这一解释的优点的最好办法,是找到另一种办法,以此来确定人们是否认为总裁的行为是故意的——这个办法并不依赖于涉及人们对“故意的”一词的使用的语用学考虑。我们可以拿用第二种办法得到的结论来与最初的结论作比较。如果用第二种办法得到的结论明显区别于最初的结论,那么这将增加亚当斯与斯特德曼的观点的可信度——他们认为最初的结论源自于语用学考虑,且没有告诉我们任何关于我们的“**故意行为**”的大众概念的本性的东西。一方面,如果用第二种办法获得的结论与最初的结论一致,那么这将增加如下这一看法的可信度,即最初的结论**并不**源自于语用学考虑,且确实告诉了我们一些关于我们的“**故意行为**”的大众概念的本性的东西。

诺布(2004 年 b)问人们这个总裁是不是**为了**增加利润而改善或破坏环境的。根据诺布的看法,人们并不愿意说某人是为了达到某个具体目标而执行某个行为的,除非他们相信她是故意做出这个行为的。这样一来,问人们这个总裁是不是为了增加利润而改善或破坏环境这一点就为我们提供了一个间接的办法,以此来确定人们是否认为这个总裁的行为是故意的。诺布

---

① 可以给出一个类似的说法来解释为什么人们倾向于说总裁不是故意地改善环境的。在**改善环境**例子中,人们不想因改善环境而赞扬这个总裁。如果他们说这个总裁是故意改善环境的,那么这将意味着他应该为他的行为而得到赞扬。结果,当人们被问及这个总裁是否故意改善环境的时候,人们倾向于说总裁并不是故意改善环境的。再一次地,人们说总裁并不是故意改善环境的倾向并没有反映他们的“**故意行为**”的大众概念的本性;它仅仅反映了他们想要克制住因改善环境而赞扬总裁的愿望。

发现,大多数被要求考虑**危害环境**例子的人都愿意说这个总裁是为了增加利润而破坏环境的,而大多数被要求考虑**改善环境**例子的人都不愿意说这个总裁是为了增加利润而改善环境的。① 由于用第二种办法获得的结论与最初的结论相符合,诺布得出结论,认为最初的结论并不源自于语用学考虑,且确实告诉了我们一些关于我们的"**故意行为**"的大众概念的本性的东西。

诺布(2004 年 b)的研究无法说服亚当斯和斯特德曼。在 56
一篇对这一项研究作出回应的文章中,亚当斯和斯特德曼(2004b)提出,在人们说这个总裁故意损害环境的意愿中扮演着某种角色的那种语用学考虑,也在人们说这个总裁是为了增加利润而损害环境的意愿中扮演着某种角色。他们同意诺布的看法,即人们不愿意说某人是为了达到某个目的而做出某个行为的,除非他们相信她是故意做出这个行为。但是,他们提出,这正是为什么我们应该认为在人们说这个总裁故意破坏环境的意愿中扮演着某种角色的语用学考虑也在人们说这个总裁是为了增加利润而破坏环境的意愿中扮演着某种角色的原因。

亚当斯和斯特德曼是对的。如果语用学考虑在人们说这个总裁故意破坏环境的意愿中扮演着某种角色,那么它们也有可能在人们说总裁是为了增加利润而改善环境的意愿中扮演着某

① 更具体地说来,用一个从－3(听上去是错的)到＋3(听上去是对的)的评分范围,那些被要求考虑**危害环境**例子的人被问及了这样一个问题,即说这个总裁是为了增加利润而损害环境听上去是不是对的,而那些被要求考虑**改善环境**例子的人被问及了这样一个问题,即说这个总裁是为了增加利润而改善环境听上去是不是对的。那些被要求考虑**危害环境**例子的人的平均给分是＋0.6 分,那些被要求考虑**改善环境**例子的人的平均给分是－1.0 分。这一差别在统计学上具有 $p=0.01$ 的显著水平。

种角色。然而，不清楚的是，语用学考虑实际上扮演的角色是不是亚当斯和斯特德曼所说的那一种。马勒(Malle)(2006年)指出，那些被要求考虑**没有技巧/不道德结果**例子的人被引导去直接表达他们的谴责。认为人们会觉得有必要在会话中暗示某些他们被允许去直接表达出来的东西，这是很奇怪的。① 情况仍然可能是，以一种不同的方式，人们对故意性的认定受到了他们对责任的认定的影响：人们作出对故意性的认定，并不是为了在会话中暗示某种责任，而是为了一致性的考虑。也就是说，人们也许会觉得，既说某人要为她的行为负责，又说她不是故意做出这些行为的，这是不一致的。但是，似乎情况也不是这样的。相反，似乎人们很愿意为非故意行为而指责他人。考虑一下如下的小场景：

喝醉了的司机

下班后，鲍伯在一个聚会中喝得烂醉。当聚会结束，他踉跄着钻进汽车，开车回家。那时他已经很醉了——以至于无法操控汽车，他与迎面而来的车子相撞，撞死了一家五口人。②

57 诺布(2003年b)发现，大多数被要求考虑这个小场景的人都判断鲍伯要为杀死这一家五口人而负责，即使他们也判断鲍

---

① 马勒(2006年)给出了他自己对这些结论的一些解释。在那些最有趣的解释中，有一个是这样的，即副作用效应最好被解释成是源自于小场景被给出的方式的。根据马勒的看法，因为参与者的注意力集中在这个小场景的可评价的部分，他们可能会认为，在作出故意性判断时，他们应该动用道德评判。在其他的背景中，他们的注意力不会如此集中，由此，他们可能不会让评价性判断来主导故意性判断。

② 诺布(2003年b)从米尔(2001年)那里借用了这个例子。

伯这样做不是故意的。因此,似乎人们作出对故意性的认定,既不是为了在会话中暗示什么责任(因为在可以直接认定责任的背景下,人们也做出了相同的认定),也不是为了一致性的考虑(因为存在这样一些背景,在这里,人们愿意在没有对故意性做出认定的情况下对责任进行认定)。将这两种考虑放在一起,似乎很难继续维持副作用效应最好用语用学考虑来解释这一看法了。

4. 副作用效应和我们的“故意行为”概念

如果语用学考虑无法恰当地解释副作用效应,诺布(2003年a,2006年)提出也许这种效应告诉了我们一些非常具体的东西,这些东西是与规范性判断和人们的“故意行为”概念的关系,以及与掌管着人们关于**故意行为**的判断的那些潜在机制息息相关的。这个观点实际上是非常激进的,它标志着对传统观点的背离,有关我们的“故意行为”概念的传统观点认为这一概念的作用局限于对行为进行解释和预测。诺布认为,我们的“故意行为”概念不仅在我们的解释和预测活动中扮演着某种角色,还在我们的**评价性**活动中扮演着某种角色。这使我们能够解释这些结论,而无须将它们解释掉。由于我们的“故意行为”概念在我们的评价性实践中扮演着某种角色,那么规范性判断影响故意性判断也是很恰当的。

托马斯·纳德尔霍弗(2004年a,2004年b,2006年b)同意诺布的看法,这些结论告诉了我们一些关于人们的道德判断与“故意行为”概念的关系的东西。然而,他反对这些结论告诉了我们一些有关掌管着我们关于故意行为的判断的潜在机制的恰

58 当功能的东西，他转而提出，它们揭示了这种机制**发生故障**的方式。①若把诺布的**危害环境**例子当作我们的典型，这里的一般看法是这样的：那些被要求考虑**危害环境**例子的人对这个总裁有着负面的情感反应；这种负面的情感反应扭曲了人们的故意性判断，使他们不正确地判断这个总裁是故意损害环境的。②

当然，要说人们负面的情感反应扭曲了他们的故意性判断，这意味着道德考虑**不应该**影响我们的故意性判断。为什么不呢？在纳德尔霍弗看来，其理由在于：某人是否故意做出某个行为这一点应该在决定她是否应该为这个行为负法律责任时扮演着某种角色。如果人们的负面的情感反应影响了他们的故意性判断，而且故意性判断决定了关于法律责任的判断，那么当某个行为产生强烈的负面情感反应时，人们将更有可能判断某人要为该行为负法律责任。为了看清这可能带来的问题，考虑一

---

① 马勒和尼尔森(Nelson)(2003 年)也提出这些结论最好用情感偏见来解释。他们研究了一些激烈争吵的夫妇(在这种争吵中，每一方都对另一方有着相当负面的情感反应)，并且发现一般说来这些夫妇会判断对方的所有行为(即使是那些在其他情况下不会被判断为故意行为的行为)都是故意的。这些研究向他们揭示了负面的情感反应会**扭曲**人们对故意性的判断。

② 纳德尔霍弗(2004 年 a)提出正面的情感反应似乎对人们的故意性判断也有着类似的影响。考虑一下如下这个小场景：

学术帮助

假设斯蒂夫和杰森是朋友，在一次论文比赛中，他们在相互竞争。杰森决定帮助斯蒂夫编辑他的文章。作为他们共同的朋友，艾伦说："难道你没有认识到，如果你帮助了斯蒂夫，那么你将降低自己赢得比赛的几率吗？"杰森回答："我知道帮助斯蒂夫会降低我得奖的几率，但是我不在乎。我只是想要帮助我的朋友！"很确定的是，由于杰森的帮助，斯蒂夫赢得了比赛。

纳德尔霍弗发现大多数被要求考虑**学术帮助**例子的人对杰森都有着正面的情感反应，并且判断杰森是故意降低他自己赢得比赛的几率的。

下如下两个小场景：

警察

设想一个小偷驾驶着一辆汽车，车上全是最近偷来的赃物。当他在等红灯的时候，一个警察走到车窗前，并且挥舞着一把手枪。当看到这个警察时，小偷加速穿过了十字路口。令人惊讶的是，在车子加速开走时这个警察抓住了车身的一边。小偷开始走"之"字形，想要逃脱——他完全知道这样做会让警察处于相当危险的境地。但是小偷不在乎，他就是想要逃走。警察很不幸，这个小偷成功地将他甩掉了。结果警察滚进了车流之中，受到了致命伤。几分钟后，他死了。

偷车贼

设想一个人在他车里等红灯。忽然，一个偷车贼靠近他的车窗，并挥舞着一把枪。当他看到这个小偷时，他感到惊慌了，加速穿过了十字路口。令人惊奇的是，在车子加速开走时这个小偷抓住了车身的一边。司机开始走"之"字形，想要逃脱——他完全知道这样做会让小偷处于相当危险的境地。但是司机不在乎，他就是想要逃走。小偷很不幸，司机成功地将他甩掉了。结果小偷滚进了车流之中，受到了致命伤。几分钟后，他死了。 59

尽管这两个小场景在结构上是一致的，纳德尔霍弗（2006年b）发现，比起那些被要求考虑**偷车贼**例子的人，那些被要求考虑**警察**例子的人对小车司机有着更加强烈的负

面情感反应。[①] 此外，比起那些被要求去考虑**偷车贼**例子的人，那些被要求去考虑**警察**例子的人更有可能判断司机是故意导致他人死亡的。[②] 这意味着比起偷车贼的死，人们更有可能认为那个司机要为警察的死负法律责任。但是，在纳德尔霍弗看来，我们并不希望发生这样的事情，一个人被指控犯下某种罪行，若这种罪行产生了强烈的负面情感反应，那么人们就更有可能认为她要为这种罪行负责。这样说来，纳德尔霍弗极力主张道德考虑**不应该**影响我们的故意性判断。道德考虑对我们的故意性判断的影响指示着掌管着我们关于故意行为的判断的那种潜在机制的某种故障。

让我们来更为仔细地审查这个观点。这里的核心看法是人们的情感反应在主导着(在这个例子中，这种主导是轻率的)他们的故意性判断。人们觉得这个总裁需要为破坏环境负责，这种负面的情感反应导致他们(不正确地)判断总裁是故意这样做的。问题在于故意性判断并不总是情感性的。实际上，我们发现即使是那些对总裁的行为无法作出**任何**情感反应的人(比如那些脑内前额叶皮层损害[ventromedial prefrontal cortex damage]的人)也会判断这个总裁是故意破坏环境的(杨[Young]

---

① 更具体地说来，用一个从0(没有责任)到6(有很大的责任)的评分范围，被要求考虑**警察**例子的人被问道这个小偷应该为警察之死负多大的责任，被要求考虑**偷车贼**例子的人被问道这个司机要为小偷的死负多大的责任。那些被要求考虑**警察**例子的人给出的平均分是5.11分，而那些被要求考虑**偷车贼**例子的人给出的平均分是2.01分。纳德尔霍弗把人们的负面情感反应和他们认定责任的意愿联系了起来，并得出结论：比起那些被要求考虑**偷车贼**例子的人，那些被要求考虑**警察**例子的人有着更为强烈的负面情感反应。

② 更具体地说来，在那些被要求去考虑**警察**例子的实验对象中，有37%的人判断这个贼故意导致警察的死亡，而在被要求去考虑**偷车贼**例子的实验对象中，仅有10%的人判断司机是故意导致贼的死亡的。这一差别在统计学上达到p值小于0.001的显著水平。

等人,2006 年)。这就质疑了如下看法:是针对总裁的行为的情感反应导致人们判断他故意损害环境的。当然,责任也并不总是情感性的。人们当然也有可能认为这个总裁需要因破坏环境而被谴责,同时无须对总裁行为具有相应的情感反应。但是,正 60
如**喝醉了的司机**例子所完美地揭示的那样,责任和故意性也可以各走一边。在某个行为被认为不是故意的情况中,人们似乎也很愿意作出责任认定。① 还有,即使在那些责任和故意性并不分离的例子中,我们发现人们一般会在他们作出责任判断**之前**做出故意性判断(古列尔莫[Guglielmo]和马勒)。这就使如下观点很难维系了,即人们之所以判断总裁是故意破坏环境的,是因为他们想要因为这样做而谴责总裁。② 不管是什么在主导着人们的故意性判断,似乎那个东西既不是**情感**也不是**责任**。③

约书亚·诺布和托马斯·纳德尔霍弗同意这些结论告诉了我们一些东西,这些东西是与人们的道德判断和他们的**故意行为**概念的关系有关的。诺布认为这些结论告诉了我们这样一些东西,这些东西是与掌管着人们关于故意行为的判断的潜在机制的恰当功能有关的,而纳德尔霍弗则认为这些结论告诉我们的是这样一些东西,这些东西是与这种机制发生的故障有关的。爱德华·马舍雷(Edouard Machery)(2008 年)质疑了这两个观点。首先,马舍雷主张我们应该对副作用效应是否最好用合格

---

① 诺布和曼德卢(2004 年)提出,这个证据还表明了在某个道德上恶劣的结果被认为是故意的例子中,人们不愿意进行责任认定。

② 实际上,古列尔莫和马勒(2010 年)提出,这一证据表明了人们经常是因为相信这个总裁是故意这样做的,所以才会因为破坏了环境而谴责这个总裁。

③ 这个观点也许太强了点。也有可能我们的故意性判断出现的时候,有很多不同的因素都在扮演着某种角色,而且**有时候**责任和情感也是这些因素的一部分。前面提到的这些考虑要做的事情,就是质疑只有这些因素掌管着我们的故意性判断这一看法。

的故意性判断来解释这一点保持中立。其次，他认为这些结论确实告诉了我们一些东西，这些东西并不与人们的道德判断和他们的**故意行为**概念之间的关系有关，而是与人们的**故意行为**概念和对损失及收益(costs and benefits)的考虑的关系有关。

在将我们的注意力转移到那个已经为人瞩目的第二个观点之前，有必要花一点时间来看看马舍雷所提倡的中立性。正如我们已经看到的，在诺布和纳德尔霍弗的争论中，其中至少有一部分问题是，副作用效应告诉我们的东西是与人们的**概念能力**还是**概念运用**有关的。能力与运用的区分要追溯到诺曼·乔姆斯基(Noam Chomsky)(1965 年)对语言能力和语言运用的区分，并且一直被用来区别一个人对某个给定概念的知识与她对这一概念的使用。这里的核心观念是一些因素(比如来源限制[resource limitations]或来自其他认知过程的干扰)可以影响一个人对某个给定概念的使用，同时不影响她关于这个概念的知
61 识或她对这个概念的意义的反思。诺布认为副作用效应告诉了我们一些关于人们的概念能力的东西，而纳德尔霍弗认为它告诉我们的东西是与人们的概念运用有关的。问题在于，任何一种对这一分歧的解决都需要一些我们目前尚不具备的东西，也就是一些公认的方法，让我们能够区分什么是、什么又不是我们的**故意行为**的大众概念的组成部分。① 诺布似乎想要让经验性证据来做这项工作，但是这种处理概念能力问题的方法的缺陷在于诺布想到的那种经验性证据并不十分适合于回答这几类问题。为了把能力和运用区分开(至少以一种能有助于解决这个特定争论的方式将能力和运用区分开来)，我们需要阐述我们的**故意行为**的大众概念应当为我们做的那些事情，然而光有经验

---

① 对此的其他讨论，可见亚历山大等人(2010 年 a，2010 年 b)。

性证据是无助于带来这样一种阐述的。

纳德尔霍弗只是规定了我们的**故意行为**的大众概念应该为我们做的那些事情,以此避开了这个问题。这使他坚持副作用效应告诉了我们一些关于人们的概念运用的东西,但是他的做法留下了这样一种可能性,即其他人(比如诺布)可以规定另一种阐述,根据这另一种阐述,副作用效应告诉了我们一些关于人们的概念能力的东西。毕竟,规定是随意的。当然,也有可能存在确定我们的**故意行为**的大众概念应该做的那些事情的其他方法,比如进化论的或目的论的进路。同时也有可能存在一些为能力与运用的区分奠基的方式,这种方式不必以我们的**故意行为**的大众概念应该做的那种事情为出发点。但是,如果这些东西缺了一种,那么似乎最合理的做法就是对于副作用效应告诉我们的是有关人们的概念能力还是概念运用的东西这一点保持中立。

结束上述讨论之后,让我们更仔细地审查一下马舍雷的第二个主张,即那些结论实际上告诉我们的东西并不与人们的道德判断和他们的**故意行为**概念的关系有关,而是与人们的**故意行为**概念和对损失及收益的考虑的关系有关。回想一下诺布(2003 年 a)的发现,大多数被要求考虑**环境危害**例子的人都判 62
断这个总裁是故意破坏环境的,同时大多数被要求考虑**改善环境**例子的人都判断总裁并不是故意改善环境的。为了解释这些结果,马舍雷提出了他称之为交易假说(trade-off hypothesis)的方案。在**危害环境**的例子中,有一种损失是与利润的增加关联着的,那就是环境的破坏。因为人们相信损失都是故意导致的,这就解释了为什么那些被要求考虑**危害环境**例子的人判断总裁是故意破坏环境的。相反,在**改善环境**的例子中,没有什么损失是与利润的增加关联着的。这解释了为什么那些被要求考虑**改**

善环境例子的人判断总裁并不是故意改善环境的。

马舍雷用两对小场景来支持他的交易假说。让我们从第一对小场景开始吧：

多掏的美元

乔感到非常口渴，因此他去了本地的一家奶茶店，要买那里卖的最大杯的饮料。在点单之前，收银员告诉他大杯奶茶现在比过去贵了一美元。乔回答："我不在乎我要多掏一美元，我想要你们这里最大杯的奶茶。"当然，乔得到了大杯奶茶，并且为此多掏了一美元。①

免费的杯子

乔感到非常口渴，因此他去了本地的一家奶茶店，要买那里卖的最大杯的饮料。在点单之前，收银员告诉他如果他买大杯奶茶，他将获得一个特别的纪念版杯子，用来装奶

---

① 费伦和萨尔吉希安(2009 年)提出**多掏的美元**例子有两个问题。首先，如果人们认为为了一杯奶茶而多付一美元是一种损失，那么他们有可能会认为这是一件坏事(尽管不是道德上的坏)。这样一来，这个例子似乎不是在对副作用的两种解释(根据其中的一种解释，人们关于某个行为是不是故意做出的判断受到了规范性考虑的影响；根据另一种解释，人们关于某个行为是不是故意做出的判断仅仅受到损失和收益的考虑的影响)之间作出决定的恰当例子。(对于针对下述交易假说的主体导向的解读，费伦和萨尔吉希安有着类似的担忧，也就是说它接近了诺布自己的假说，即人们关于某个行为是不是故意做出的直觉受到了规范性考虑的影响。)其次，**多掏的美元**例子实际上是不是包含了一种可预见的副作用，这一点是不清楚的。要注意的是，只有当一个行动者选择去做一个行为，她预见这个行为将会带来某个结果，而这个人并不是因为要带来这个结果而做出这个行为的，这个结果才是可预见的副作用。但是，为了带来某个特定的结果而做出的某些事情并不是为了带来那个结果而做出的行为的副作用；它是那个为了带来这个结果而做出的行为的一部分。这样一来，由于这个行动者为了得到大杯奶茶而多付了一美元，因此多付一美元并不是副作用。

茶。乔回答:"我不在乎什么纪念版杯子,我想要你这里最大杯的奶茶。"当然,乔得到了装在纪念版杯子里的大杯奶茶。

马舍雷发现大多数被要求考虑**多掏的美元**例子的人都判断乔是故意多付一美元的,同时大多数被要求考虑**免费的杯子**例子的人都判断乔并不是故意买纪念版杯子的。① 由于大多数人认为这种行为(多付一美元和买一个纪念版杯子)在道德上是中 63
性的,这为如下看法提供了证据,即某些不是人们关于某个行为(或该行为的结果)的道德状态的信念的东西影响了他们关于某个行为是不是故意做出的直觉。这反而成为了人们关于某个行为是不是故意做出的直觉受到损失和收益的简单考虑的影响的证据。

损失和收益的简单考虑对人们的故意性判断的显著影响不仅仅局限于涉及道德上中立的行为的例子。同样的影响似乎还可以在涉及道德上**恰当的**(morally appropriate)行为的例子中被发现。为了看清这一点,让我们考虑一下马舍雷的第二对小场景:

工人

约翰站在一条电车轨道边上。他发现电车的刹车坏了。在那条轨道上,有五名工人正在背身工作。约翰看到这失控的电车正朝着这五个人开去,如果电车按当前的路

---

① 更具体地说来,那些被要求去评价乔在**多掏的美元**的例子中的行为的人中,有95%的人判断他是故意多付一美元的,同时,在那些被要求去评价乔在**免费的杯子**的例子中的行为的人中,只有45%的人判断他是故意买这个纪念版杯子的。这一差别在统计学上达到p值小于0.001的显著水平。

> 线往前开，这五个人将会被撞死。挽救这五个人的惟一办法就是按一个开关，这个开关会把电车转到岔道上去。不幸的是，岔道上也有一名工人在背身工作。约翰知道如果他按这个开关，岔道上的这名工人将会被撞死，但是那五个人的生命就会被挽救。约翰决定按这个开关。当然，这列电车转到了岔道上，主干道上的这五个工人获救了，岔道上的那个工人被撞死了。

狗

> 约翰站在一条电车轨道边上。他发现电车的刹车坏了。在那条轨道上，有五名工人正在背身工作。约翰看到这列失控的电车正朝着这五个人开去，如果电车按当前的路线往前开，这五个人将会被撞死。挽救这五个人的惟一办法就是按一个开关，这个开关会把电车转到岔道上去。还有，轨道上还有一条背身的狗。约翰知道如果他按这个开关，岔道上的这五个工人和这条狗都将被挽救。约翰想："我不在乎是不是要拯救这条狗。我只是想要拯救这五个
> 64 人。"约翰决定按这个开关。当然，这列电车转到了岔道上，主干道上的这五个工人和狗获救了。

马舍雷发现大多数被要求考虑**工人**例子的人都判断约翰是故意导致岔道上那名工人的死亡的，大多数被要求考虑**狗**例子的人都判断约翰并不是故意挽救这条狗的。① 因为大多数人认

---

① 更具体地说来，在那些被要求评价约翰在**工人**的例子中的行为的实验对象中，有56%的人判断他是故意导致岔道上工人的死亡的，同时，在那些被要求评价约翰在**狗**的例子中的行为的实验对象中，只有23%的人认为约翰是故意拯救这条狗的。这一差别在统计学上达到p值小于0.01的显著水平。

为这个行为(导致岔道上那名工人的死亡和拯救这条狗)是道德上恰当的行为,这一点为如下观点提供了额外的证据,即某种不同于人们关于一个行为(或该行为的结果)的道德状态的信念的东西影响了他们关于这个行为是不是故意做出的直觉。[①] 再一次地,似乎人们关于某个行为是不是故意做出的直觉受到了损失和收益的简单考虑的影响。

让我们更为仔细地查看一下马舍雷的交易假说,根据这一假说,人们关于某个行为是不是故意做出的道德直觉受到了损失和收益的简单考虑的影响。对于这一假说,有两种可能的解读。[②] 根据第一种解读,这个虚构的行动者将某物视为一种损失,这影响了人们关于某个行为是不是故意做出的直觉。让我们把这种解读称为**行动者导向的解读**。按照第二种解读,人们将某物视为一种损失,这影响了人们关于某个行为是不是故意做出的直觉。让我们把这种解读称为**主体导向的解读**。

困难在于,不管是哪种解读,这个交易假说似乎都存在着严重的问题。让我们先从行动者导向的解读开始吧。罗纳德·马伦(Ronald Mallon)(2009 年)指出,如果人们关于某个行为是不是故意做出的直觉受到了这个虚构的行动者将某物视为一种损失的影响,那么我们应该预测,在如下这类例子中这种副作用效应该会消失,在这类例子中,那些虚构的行动者并不把相关的副作用视为一种损失。但是,考虑一下如下这一对小场景:

---

① 费伦和萨尔吉希安(2009 年)指出,用这一项研究来质疑诺布的假说,这里存在着一个问题。大多数人都认为这两种行为是道德上恰当的,费伦和萨尔吉希安猜想大多数人会认为第一个行为(导致岔道上工人死亡)是坏的(尽管也许不是道德上坏的),而第二个行为(拯救那条狗)是好的(也许在道德上是好的)。如果这是对的,那么实际上诺布的假说将会预测出这些结果。

② 见马伦(2009 年)及费伦和萨尔吉希安(2009 年)。

有害的恐怖分子

在一伙恐怖分子中，有一个人走到首领前面说："我们
65 正在策划炸掉一家夜总会。这将杀死很多美国人，但是它也会损害到澳大利亚人，因为很多澳大利亚人也会被杀死。"首领回答："我承认伤害澳大利亚人是一件好事，但是我不在乎。我只是想尽可能多地杀死美国人！让我们炸掉这家夜总会吧！"他们确实炸掉了这家夜总会，当然，澳大利亚人受到了伤害，因为很多澳大利亚人被杀死了。

有益的恐怖分子

在一伙恐怖分子中，有一个人走到首领前面说："我们正在策划炸掉一家夜总会。这将杀死很多美国人，但是它会降低地价，帮助周围的孤儿院得到孩子们所需要的土地。"首领回答："我承认帮助孤儿院是一件好事，但是我不在乎。我只是想尽可能多地杀死美国人！让我们炸掉这家夜总会吧！"他们确实炸掉了这家夜总会，当然，由于地产价格的下降，孤儿院得到了帮助。

马伦(2009年)发现，大多数被要求考虑**有害的恐怖分子**例子的人都判断这帮恐怖分子是故意伤害澳大利亚人的，而大多数被要求考虑**有益的恐怖分子**例子的人都判断这帮恐怖分子并不是故意帮助孤儿院的。① 这为**反对**如下观点——即人们关于

---

① 更具体地说来，在那些被要求评价恐怖分子在**有害的恐怖分子**例子中的行为的实验对象中，有92%的人判断这帮恐怖分子是故意伤害澳大利亚人的，而在那些被要求评价恐怖分子在**有益的恐怖分子**例子中的行为的实验对象中，只有12%的人判断这帮恐怖分子是故意帮助这个孤儿院的。这一差别在统计学上达到p值小于0.000 1的显著水平。马伦(2009年)将这些小场景建立在诺布和凯利(2009年)引进的例子之上，马伦发现，从有着相同基本结构的另外一对小场景中也可以得到类似的结论。

某个行为是不是故意做出的直觉受到了虚构的行动者将某物视为一种损失的影响——提供了证据。一定是其他的某些东西在影响人们关于这个行为是不是故意做出的直觉。

也许主体导向的解读是正确的。马克·费伦和哈格普·萨尔吉希安(Hagop Sarkissian)(2009 年)指出，如果人们关于某个行为是不是故意做出的直觉受到了他们自己将某物视为一种损失的影响，那么我们应当预测大多数人将会愿意判断，当他们认为某个行为是在达到某个重要目的的过程中发生的损失时，这个行为就是故意的。但是，考虑一下如下的小场景：

关切的中尉

一个中尉正在对一个中士说话。中尉给出了命令："把你的班派到汤普森高地上去。"中士说："但是，如果我把我 66
的班派到汤普森高地上，我们将把士兵直接放在敌人的火线内。他们中间的一些人肯定要被杀死!"中尉回答："听着，我知道他们将会处在火线内，我也知道他们中的一些人会被杀死。我比任何人都更关心我的士兵。但是对于这次战役的胜利来说，拿下汤普森高地是必须的。"这个班被派到了汤普森高地上面。不出意料，士兵们进入了敌人的火线，其中的一些人被杀死了。

费伦和萨尔吉希安发现，大多数被要求去考虑**关切的中尉**例子的人都判断这个中尉并不是故意导致士兵死亡的。由于人们一般会认为士兵的死亡是一种损失，这就为反对如下观点——即人们关于某个行为是不是故意做出的直觉受到了他们自己将某物视为一种损失的影响——提供

了证据。[1][2] 再一次地，肯定有其他某些东西在影响着人们关于某个行为是不是故意做出的直觉。

到目前为止，我们一直在考察对于副作用效应的这样一种解释，根据这种解释，这些结论告诉了我们一些非常具体的东西，这些东西是与规范性判断和人们的**故意行为**概念的关系有关的。所有这些解释都假定了只有一种**故意行为**的大众概念。肖恩·尼科尔斯和约瑟夫·乌拉托夫斯基(2007 年)质疑了这一假设，他们提出有很多个**故意行为**的大众概念，并且提出副作用效应最好这样来解释，即它们告诉了我们一些东西，这些东西是与规范性判断和这些不同的**故意行为**的大众概念间的关系有关的。

有争议的是，怎样才能最好地解释那些占少部分的人的回应。大多数被要求考虑**环境危害**例子的人都判断这个总裁是故意破坏环境的，同时大多数被要求考虑**改善环境**例子的人都判断这个总裁不是故意改善环境的，然而这些回应并不是单一的。一些被要求考虑**危害环境**例子的人判断总裁

---

① 由于这个小场景指明了这个中尉也认为士兵的死亡是一种损失，所以这一结果为反对如下观点——即人们关于某个行为是不是故意做出的直觉受到了这个虚构的行动者将某物视为一种损失的影响——提供了额外的证据。

② 费伦和萨尔吉希安(2009 年)有另一个反对主体导向解读的论证。他们认为，如果主体导向解读是正确的，那么我们应该预测，比起将某个行为视为在达到一个不重要的目标的过程中产生的损失，当人们将某个行为视为在达到一个重要目标的过程中产生的损失时，人们将更愿意判断一个行为是故意的。因为他发现经验证据表明了情况并不是这样的，所以他们得出结论，认为主体导向的解读是不正确的。问题在于，根据交易假说的任何一个版本，我们都没有理由预测比起将某个行为视为在达到一个不重要的目标的过程中产生的损失，当人们将某个行为视为在达到一个重要目标的过程中产生的损失时，人们将更愿意判断一个行为是故意的。损失就是损失。对于交易假说而言，没有什么东西取决于这个行动者导致这种损失是不是明智的。这里的重要的东西只是那种被引起的损失是故意引起的。

并不是故意破坏环境的，而一些被要求考虑*改善环境*例子的人判断总裁是故意改善环境的。我们该如何解释这少部分人的回应呢？

一般说来，这样的回应被视为噪音，这种噪音源自于某种执行错误。尼科尔斯和乌拉托夫斯基(2007 年)给出了一个不同 67
的解释。他们要求人们同时考虑*危害环境*和*改善环境*的例子。他们发现了一个有趣的回应模式。一些被要求考虑这两个例子的人给出了不对称的回应，他们判断这个总裁是故意破坏环境但不是故意改善环境的。其他被要求考虑这两个例子的人给出了对称的回应，他们的判断要么是这个总裁既故意破坏环境又故意改善环境，要么是这个总裁既不是故意地破坏环境又不是故意地改善环境。此外，他们还发现，那些判断这个总裁故意破坏环境的人一般都会用总裁的环境将会被破坏的*信念*来解释他们的判断，同时那些判断这个总裁并不是故意改善环境的人一般都会用总裁缺乏改善环境的*愿望*来解释他们的判断。这种回应模式，以及相应的解释，导致尼科尔斯和乌拉托夫斯基提出实际上有两个彼此无关的*故意行为*的大众概念，一个涉及一种认知的成分，另一个涉及一种动机的成分。

费里·库什曼(Fiery Cushman)和阿尔弗雷德·米尔(Alfred Mele)(2008 年)指出了尼科尔斯和乌拉托夫斯基的观点(存在两种截然不同的*故意行为*概念，一个涉及一种认知的成分，另一个涉及一种动机的成分)的一个问题。在*危害环境*和*改善环境*的例子中，总裁有一个关于这个增加利润的投机会对环境带来何种影响的信念，但是缺乏相应的损害及改善环境的愿望。库什曼和米尔想要知道的是，人们对那种行动者有着相应的愿望然而缺乏相应的信念的例子，会有什么样的回应。库什曼和米尔预测，如果尼科尔斯和乌拉托夫斯基是对的，那么这里

应该发生了三件事情。首先,比起行动者有着相应的愿望然而缺乏相应的信念的情况,在那些行动者有着相应的信念但是缺乏相应的愿望的情况下,那些动用认知性概念的人应该更有可能判断这个行动者是故意这样做的。其次,比起行动者有着相
68 应的信念然而缺乏相应的愿望的情况,在那些行动者有着相应的愿望而缺乏相应的信念的情况下,那些动用动机性概念的人应该更有可能判断这个行动者是故意这样做的。再次,比起那些动用认知性概念的人,那些动用动机性概念的人应该更有可能判断一个行动者在有着相应的愿望但缺乏相应的信念的情况下是故意这样做的。与这些预测相反,库什曼和米尔发现,那些动用动机性概念的人并不比那些动用认知性概念的人更有可能判断这个行动者在具有相应愿望但缺乏相应信念的情况下是故意这样做的,那些动用认知性概念的人不会判断这个行动者在具有相应信念的情况下比在具有相应愿望然而缺乏相应信念的情况下更有可能是故意这样做的。库什曼和米尔得出结论,认为实际上存在着多于两种的不同的*故意行为*概念:一种把愿望视为故意行为的必要条件;另一种把愿望视为故意行为的充分条件;还有一种,*只有在道德上为恶的行为的例子中*,才把愿望视为故意行为的充分条件。

不管这里有多少种概念,多种概念解释的优点在于它解释了少部分人的回应,而无须将它们解释掉。它还为副作用效应本身提供了一种简洁的解释。在*危害环境*的例子中,这个总裁明知破坏了环境,但是他并不在乎这一点;在*改善环境*的例子中,这个总裁明知改善了环境,但是他并不在乎这一点。大多数被要求考虑*危害环境*例子的人都判断这个总裁是故意破坏环境的,而大多数被要求考虑*改善环境*例子的人都判断这个总裁并不是故意改善环境的,解释这一事实的东西是不同的规范性判

断触发了不同的**故意行为**概念。

5. 结　论

似乎副作用效应告诉了我们**一些东西**,这些东西是与规范
性考虑和我们的**故意行为**概念(或几个**故意行为**概念)的关系有
关的。迪恩・佩蒂特(Dean Pettit)和约书亚・诺布(2009 年)
提出这种影响的范围还要更广一些,他们认为,它告诉我们的东 69
西是关于规范性考虑和**一般说来**的大众心理的关系的。他们发
现规范性考虑广泛影响了一系列大众心理判断,包括关于决定、
愿望、喜好、支持、选择的判断。最近,其他研究者又做了一些研
究,表明规范性考虑影响了大众关于因果性(希区柯克
[Hitchcock]和诺布,2009 年,洛克希布劳[Roxborough]和坎比
[Cumby],2009 年)和知识(毕比[Beebe]和布克瓦尔特,2010
年)的心理判断。随着更多研究的陆续出现,似乎我们现在才开
始理解这种影响会是多么广泛。

这种副作用效应实际上告诉我们的是什么,这尚待确定。具体说来,它告诉我们的东西是关于我们的概念能力还是关于概念运用的,这是尚待确定的。在这里,实验哲学家面临一个难题。要确定副作用效应是否告诉了我们一些关于我们的概念能力的东西,这需要一些实验的和理论的资源,而实验哲学家目前并没有这类资源。这里所需要的,是对我们的大众概念所应该做的工作的考察,或者是对相关认知过程与机制的物理执行的考察。不幸的是,实验哲学家一般采用的实验方法,即调查的方法,并不适于产生这两类考察的任何一种。我们需要的是对掌管着我们的大众心理判断的认知过程、机制的神经解剖学的考察,以及对我们的大众概念所应该做的

工作的进化论的(或目的论的)考察。要发展这种考察,需要实验哲学家拓展他们所采用的各种实验方法,成为更加具有实验精神的哲学家。

第四章
实验哲学和哲学方法论

1. 引　　言

我们的哲学直觉为我们提供了关于这个世界和我们自身的 70
重要洞见这一看法并不是没有遭到过批评，其中大部分批评都集中在哲学直觉在关于哲学理论的真或表面合理性的争论中的所谓证据地位上。一些人提出，可以把哲学直觉视为某物的证据，但它们不能被视为哲学家真正应该感兴趣的东西的证据（比如，可见科恩布利思，1998 年、2002 年、2007 年）。其他一些人提出，哲学直觉不能被视为证据，因为我们不能事先确定它们是不是真理的可靠向导（比如，可见卡明斯[Cummins]，1998 年）。还有其他一些人提出哲学直觉不能被视为证据，因为它们是可错的（比如，德维特[Devitt]，1994 年；埃尔金[Elgin]，1996 年）。这些质疑是富有启发的，但是，相比它们告诉我们的那些关于哲学对我们的哲学直觉的兴趣的地位的东西，这些质疑之所以富有启发，更多地是因为它们告诉我们的那些关于一个成功的质疑看上去应该是怎么样的东西。我们从中学会了一个成功的质疑不能依赖于一个过于狭隘的哲学观念，不能苛求那些认知上不可能的东西，也不能太强硬，以至于认为*所有*假定的证据都是 71
不可信的。

在本章中,我们将关注另一种不同的质疑——这种质疑是由实验哲学家提出的,还要询问一下这种质疑对前面那些经验教训吸取得怎么样了。哲学家热衷于用我们的哲学直觉的镜头来捕捉哲学问题,这一做法至少部分地源自于他们相信这些直觉或多或少是被普遍享有的。然而,近来的经验性工作发现了直觉多样性的一些有趣的形式。在方法论上,直觉多样性给出了自己的质疑,这迫使我们要么将这些证据解释掉,要么将我们关于这些证据多样性的担忧解释掉,然而真正的问题在于这些直觉多样性的形式表明了我们的哲学直觉可能敏感于这样的一些东西,我们从未预料到而且也不希望这些东西是这样的。由于不受欢迎的证据敏感性只有在它可以被预测和控制的情况下才是可以接受的,因此似乎基于直觉的哲学实践面临着一个质疑。

这一质疑通常被称为**限定主义质疑**(restrictionist challenge),它的确切性质,以及应该从这项工作中得出一些什么样的普遍的方法论结论,被证明是很难确定的。这一质疑有一个最激进的版本,它要求在方法论上彻底消除哲学直觉,但是这一立场似乎太激进了,它既没有得到经验结论的支持,同时,为了顺应那些结论,它也不是必需的。这种质疑的最保守版本要求在方法论上作一些有限的限定,移除那些成问题地敏感的哲学直觉(problematically sensitive philosophical intuitions),保持我们基于直觉的实践不受影响。这个立场似乎太保守了,它无法认识到如下这一点的危险性,即不知道这种成问题的直觉敏感性有多么普遍。正确的立场介于两者之间,它要将一种**局部的**方法论限定和我们看待和处理基于直觉的实践的方式的**全面**转变结合在一起。真正的质疑并不仅仅在于直觉并不是完全可靠的这一事实,还在于实际上我们对它们所知甚少。

我们缺乏解释那种成问题的直觉敏感性的方法，由此也缺乏搞清楚它的规模，确定补偿这种敏感性的策略，预测它在何处会出现的方法。我们真正需要的，是对哲学直觉的一种普遍而系统的理解。通过更好地理解它们是什么，它们来自哪里，72
什么因素在影响着它们，我们可以更好地理解它们在哲学实践中能够扮演何种角色。

2. 直觉多样性

我们倾向于相信我们的哲学直觉差不多是被普遍享有的。这似乎部分地为我们基于直觉的实践提供了保证。我们认为这些实践不仅仅是我们的思想自传(intellectual autobiography)，我们诉诸哲学直觉，我们这样做，是因为我们预计到了其他人也有和我们一样的直觉判断。但是，最近的经验研究表明我们可能是错的：我们所发现的不是普遍性，而是多样性。

有一种被广泛探讨的直觉多样性，那就是**文化的**多样性：我们的哲学直觉似乎敏感于我们的文化背景。① 斯蒂芬·斯蒂克(Stephen Stich)(1988 年，1990 年)是第一个严肃提出文化背景可能在我们的哲学直觉的塑造中扮演着某种角色的人。但是，直到他在一系列关于我们的认知直觉的跨文化研究中与乔纳森·温伯格和肖恩·尼科尔斯(温伯格等人，2001 年，尼科尔斯等人，2003 年)展开合作后，这个看法才开始扎下了根。这些

---

① 对其他形式的文化认知多样性的精彩讨论，可见尼斯比特等人(2001 年)，尼斯比特和诺伦泽杨(Norenzayan)(2002 年)，尼斯比特(2003 年)。尼斯比特及其同事发现了一系列不同的认知过程，包括直觉和记忆，也呈现出了文化的敏感性。早期对道德直觉和文化多样性的一项富有影响的研究，可见海德(Haidt)等人(1993 年)。

跨文化研究揭示了,实际上,来自不同文化背景的人们有着相当不同的认知直觉。考虑一下如下的小场景:

> **美国汽车**
>
> 鲍伯有一个朋友,名叫吉尔,她已经开一辆别克车很多年了。因此,鲍伯认为吉尔开着一辆美国车。然而,他并不知道她的别克最近被偷了,他也不知道吉尔用一辆庞蒂克(Pontiac)来代替了别克,庞蒂克是另一种美国车。鲍伯知道吉尔开一辆美国车吗?或者他只是相信这一点?

如果文化背景在塑造我们的哲学直觉时没有扮演任何角色,那么我们应该预测来自不同文化背景的人会形成关于这个例子的类似的直觉判断。令人注意的是,这并不是我们发现的
73 东西。相反,我们发现大多数来自西方背景的人都判断鲍伯只是相信吉尔开一辆美国车,大多数来自东亚或南亚背景的人都判断鲍伯确实知道吉尔开一辆美国车。① 似乎我们关于这个例子的直觉越出了我们可能作出的预计,并且,文化背景在产生这些直觉时扮演着**某种**角色。

尽管我们并不明确知道文化背景在产生关于这一例子的哲学直觉时所扮演的角色到底是**什么**,我们确实知道这个例子并不是惟一的。文化多样性不仅不局限于我们关于这一假设的具体例子的认知直觉,而且这种多样性根本不局限于认知直觉。最近的实验哲学研究表明我们的一些语义直觉也具有文化多样

---

① 具体说来,他们发现有26%的西方的实验对象,56%的东亚的实验对象和61%的南亚的实验对象判断鲍伯知道吉尔开一辆美国车。西方的实验对象和东亚、南亚的实验对象的差别在统计学上达到p值小于0.01的显著水平。

性(马舍雷等人,2004 年,马伦等人,2009 年)。比如,考虑一下克里普克那个著名的哥德尔案例的两个版本:

哥德尔

假设约翰在大学里学到哥德尔是那个证明了一条重要数学定理的人,我们称这条定理为算术的不完备定理。约翰精通数学,并且能够准确地陈述不完备定理,他认为哥德尔是该定理的发现者。但是这是他关于哥德尔所知道的惟一的东西。现在假设哥德尔并不是这一定理的作者。实际上,是一个被称为“施密特”的人做了这项工作,在多年前的维也纳,他的尸体被离奇地发现了。他的朋友哥德尔以某种方式取得了这份手稿,此后这份手稿就被认为是哥德尔的了。这样一来,他就被认为是那个证明了算术的不完备定理的人。大多数听到“哥德尔”这个名字的人都像约翰一样:哥德尔发现了不完备定理是他们关于哥德尔所知道的惟一的东西。

祖冲之

艾薇是香港的一名中学生。在天文学课堂上,她被告
知祖冲之是最早确定夏至和冬至的准确时间的人。但是, 74
像她的同班同学一样,这是她关于祖冲之所知道的惟一的东西。现在假设实际上祖冲之并没有做出这项发现,他是从一个天文学家那里偷来的,而这个天文学家在做出这项发现后很快就死了。但是,这件盗窃案完全没有被人察觉,祖冲之以发现夏至、冬至的准确时间而成名。很多人都像艾薇一样:祖冲之确定了夏至、冬至的时间是他们关于他所知道的惟一的东西。

要记住，我们关于类似这些例子的语义直觉本应该帮助我们在指称的摹状词理论和因果-历史性理论之间作出抉择。有趣的是，我们发现，比起有西方背景的人，那些有东亚背景的人更倾向于是摹状词主义者（相信约翰指的是施密特）。[①] 这表明，虽然大多数分析哲学家都追随克里普克，接受了一种指称的因果-历史性阐释，但是对于西方文化来说，这种做法可能也是有些独特的。[②]

似乎斯蒂克是对的：文化背景至少在塑造我们的一些哲学直觉时扮演着某种角色。但是，故事并未在这里终止。并非所有的直觉多样性都是文化多样性，最近，人们的关注点已经转移到了另一种人与人之间的直觉多样性，即**性别**多样性（赞姆佐[Zamzow]和尼科尔斯，2009 年，斯蒂克和布克瓦尔特，2011 年，布克瓦尔特和斯蒂克，即将出版）。考虑一下著名的**电车**例子的两个版本：

陌生人

你每天都要散步，现在你正沿着铁路散步，你注意到渐渐靠近的火车失控了。你看到了这里发生的事情：火车司

---

① 具体说来，关于这些小场景的摹状词主义判断得 0 分，而因果-历史性判断得 1 分。每个小场景的分数被加在了一起，使加起来的分数能处于 0 到 2 之间。用这个程序，马舍雷和他的同事们发现西方的实验对象比东亚的实验对象更有可能具有因果-历史性的直觉。西方的实验对象给出的平均分数是 1.13 分，东亚的实验对象则是 0.63 分，二者的区别在统计学上达到 p 值小于 0.05 的显著水平。最近对这项研究的讨论，包括理论性的和经验性的关注，可见路德维希（2007 年），杜尔歇（Deutsch）（2009 年），马蒂（Marti）（2009 年），斯特斯马和列文古德（Livengood）（2011 年）以及市川等人（即将出版）。

② 对这些经验结论的仔细考察可以表明，即使在西方文化的**内部**，这种做法也是相当独特的，因为关于这些例子，即使有西方背景的很大一部分人也有着与指称的因果-历史性阐述不一致的直觉。

机看到五个人在铁路上工作，他猛踩刹车，但是刹车坏了，司机昏了过去。火车正冲向这五个人。火车开得如此之快，以至于他们无法及时离开铁路了。你刚好站在一个开关旁边，你认识到拯救这五个人的惟一办法就是拧这个开关，它将把这列火车转到岔道上去，因此能阻止它撞死那五个人。然而，岔道上有一个陌生人背身站着，如果你拧下这个开关，那五个人将得到拯救，但是岔道上的那个人将会被撞死。

孩子

你每天都要散步，现在你正沿着铁路散步，注意到渐渐 75
靠近的火车失控了。你看到了这里发生的事情：火车司机看到五个人在铁路上工作，他猛踩刹车，但是刹车坏了，司机昏了过去。火车正冲向这五个人。火车开得如此之快，以至于他们无法及时离开铁路了。你刚好站在一个开关旁边，你认识到拯救这五个人的惟一办法就是拧这个开关，它将把这列火车转到岔道上去，因此能阻止它撞死那五个人。然而，岔道上背身站着一个 12 岁的男孩，如果你拧下这个开关，那五个人将得到拯救，但是岔道上的那个男孩将会被撞死。

我们可以预测，关于这些例子中的道德上恰当的做法是什么，男人和女人有着大致相同的直觉。结果我们错了。珍妮弗·赞姆佐(Jennifer Zamzow)和肖恩·尼科尔斯(2009 年)发现，比起男人，女人更不可能认为在**孩子**的例子中改变车子的方向是道德上可取的，而比起女人，男人更不可能认为在**陌生人**的

例子中改变车子的方向是道德上可取的。① 当我们从伦理直觉转移到认知直觉和形而上学直觉之上时，这种性别的多样性就变得更为明显了。比如，韦斯利·布克瓦尔特和斯蒂芬·斯蒂克(即将出版)描述了一项有趣的研究，这项研究是克里斯蒂安·斯塔曼斯(Christina Starmans)和奥利·弗里德曼(Ori Friedman)(即将出版)所进行的。斯塔曼斯和弗里德曼要求人们考虑如下的小场景：

> **手表**
>
> 彼得正在关着门的公寓里看书。他决定洗个澡。他把书放在茶几上。他取下手表，也放在茶几上。然后，他走进了浴室。彼得开始洗澡的时候，一个贼悄悄地潜入了彼得的公寓。这个贼偷了彼得的手表，在原来的位置上放了一块便宜的塑料手表，然后走了。彼得只洗了两分钟的澡，而且他什么也没有听到。

76 结果是令人震惊的：比起男人，女人更有可能认为彼得知

---

① 更具体地说来，实验对象被要求对他们赞同或反对这两个例子中改变电车的方向在道德上可取的程度进行评分(用一个7分制的李克特量表法，1分＝强烈反对，7分＝强烈赞同)。在**陌生人**例子中，男人的平均给分是4.21分，女人则是4.95分。在**孩子**例子中，女人的平均分是4.26分，而男人则是4.87分。男人和女人对这两个例子的回应的差别在统计学上接近 $p=0.07$ 的显著水平。除这两个例子外，赞姆佐和尼科尔斯要求实验对象考虑另一些例子，在这些例子中，岔路上的那个人是参与者的兄弟或姐妹。有趣的是，他们发现比起女人，男人判断杀死某人的兄弟在道德上更不能被接受(男人的平均分是3.41分，女人是4.33分)，同时比起男人，女人判断杀死某人的姐妹在道德上更不能被接受(女人的平均分是3.78分，男人是4.40分)。正如赞姆佐和尼科尔斯所写到的："似乎存在着一种偏袒某人自己的性别的偏见，至少在涉及兄弟姐妹和高速行驶的火车时是这样的。"

道桌子上有一块手表。① 当我们考虑我们的一些形而上学直觉时,也可以得到类似的结论。在布克瓦尔特和斯蒂克报道的另一项研究中(即将出版),吉奥弗里·霍尔兹曼(Geoffrey Holtzman)要求人们考虑如下的小场景:

坚定的杀手

假设科学家搞清楚了大爆炸时期宇宙的准确状态,并且搞清楚了所有的物理法则。他们将这一信息输入电脑,电脑完美地预测了发生过的所有事情。换言之,他们证明了,由于物理的法则和所有先前发生的事情,所有发生的事情都不得不像它们发生的那样发生。

然后吉奥弗里·霍尔兹曼问别人,这个世界中的某个人是否可能自由地选择去杀或者不杀另一个人,他发现比起男人,女人更有可能认为这是可能的。② 尽管未来的研究才是结论性的,但是似乎人与人之间的直觉多样性不仅存在于文化之**间**,还存在于文化之**内**。③

---

① 更具体地说来,只有41%的男性实验对象说彼得"确实知道"桌子上有一块手表,而71%的女性参与者说彼得"确实知道"桌子上有一块手表。这一差别在统计学上达到p值小于0.05的显著水平。为了排除这样一种可能性,即故事主角的性别主导着结果,斯塔曼斯和弗里德曼进行了另一项研究,使用了另一个稍微不同的小场景,这个小场景的主角是女性。总之,结果更加令人震惊。在这个版本中,只有36%的男性参与者认为这个女主角"确实知道",而有75%的女性参与者说她"确实知道"。这一差别在统计学上达到p值小于0.01的显著水平。

② 更具体地说来,只有35%的男性参与者说一个处在这一世界中的人能够自由地选择是否谋杀另一个人,而有63%的女性参与者说一个处在这一世界中的人能够自由地选择是否谋杀另一个人。这一差别在统计学上达到p值小于0.000 5的显著水平。

③ 事实上,有证据表明人与人之间的直觉多样性比这还要精细一些,它们遵循的是个体的人格类型。可见科克雷和菲尔兹(2009年)以及菲尔兹和科克雷(2007年,2009年)。

直觉多样性对我们基于直觉的实践提出了重大的质疑。任何一个人,只要他想从那些在回应某个假定的具体例子时产生的哲学直觉中选出一种,那么他就不得不解释为什么其他的直觉要被忽视。在某些例子中,我们不难发现看似合理的解释。正如戈德曼(2007 年)指出的,直觉有好几种搞错的方式:我们可能会对手边的例子有错误的了解,或者对它们有不充分的了解;我们可能会忘记或忽略关于这些例子的相关细节;我们关于这些例子的直觉判断可能会受到我们的理论预设的蒙蔽。问题在于,当我们面对的是小规模的直觉分歧时,这种做法也许是合理的,但是在那些有着大规模的、系统的直觉多样性的例子中,这种做法就不再是合理的了。至少我很难想象这样的解释会有什么用。(在下一章中,我们将会考虑一个相当普遍的解释,这个解释为谁的直觉可以被当作证据添加了一些限制。)

77 另一个选择,是将我们关于直觉多样性的担忧解释掉。这样做的一种方法是跟随戈德曼(2007 年),他提出,我们应该认为哲学研究的目的是帮助我们更好地理解我们个人的或共有的概念。[①] 正如戈德曼所言(2007 年,第 13 页):

---

① 赞姆佐和尼科尔斯(2009 年)给出了另一个不应该对直觉多样性如此担忧的理由。他们利用了直觉证据和经验证据的类比,提出对科学史的借鉴以及社会科学和认知科学的著作都表明了证据多样性实际上是一件认知上的好事,我们要促进它,而不是害怕它。这里的看法是证据多样性在认知进步中扮演着某种角色。我们应该把更多的注意力放在证据多样性带来的认知好处上,在这一点上,赞姆佐和尼科尔斯无疑是对的,但是就我们的目的而言,重要的是要记住改变会涉及到选择——对分歧的解答,或至少是对分歧的消除——任何一个进步的主张都需要一些将这些选择作为正确的选择而加以辩护的方法。然而,不清楚的是,他们的建议是不是让我们忽视了对解决证据多样性问题的原则性办法的需求,尽管它可能会让我们对能够找到这种方法的前景持更为乐观的态度。

> [这个观点]吸引人的主要地方在于它很好地处理了针对直觉的可靠性的质疑,这些质疑源自于人与人之间的直觉的多变性或冲突。如果目标概念被解释为个人心理意义上的概念,那么伯纳德的F适用于x的直觉仅仅对于**他的**F概念来说才是证据,同时埃克的F不适用x的直觉仅仅对于**她的**F概念来说才是证据。如果伯纳德直觉到某个给定的例子是知识的一个例子,而埃克直觉到它不是这样的,那么直觉之间的冲突可以被消除,因为这两种直觉都只对他们自己的个人概念而言才有证据的意义,而个人的概念是可以不同的。这可能正是温伯格所报道的那些例子所显露出来的东西。

这里的看法是直觉多样性反映了概念的多样性,而且,并非所有种类的概念多样性都必然是成问题的。但是,这种处理直觉多样性的方法把一种哲学观,即认为哲学告诉我们的是一些关于这个世界的东西,换成了另一种哲学观,即认为哲学告诉我们的是某些关于我们思考这个世界的方式的东西。这样做还会导致某种令很多哲学家厌恶的哲学相对主义。

3. 直觉的敏感性

当然,方才说的东西都不是结论性的。实际上,我们才刚刚触碰到了直觉多样性给我们基于直觉的实践带来的方法论质疑的表皮。但是这没有关系,因为不管直觉多样性给我们基于直觉的实践带来了什么困难,在这里,多样性都不是真正重要的东西。真正重要的是直觉多样性告诉我们的关于我们自身的东西,也就是说,我们的哲学直觉可能受到一些我们从未预料到的东西的影响(而且我们的哲学理论从未对这种影响作出解释)。 78

当我们把注意力从**人与人之间的**(interpersonal)直觉多样性转移到**个人内部**(intrapersonal)的直觉多样性时(换言之就是从人与人之间的直觉差别转移到人内部的直觉差别时),这一点就会变得愈加清楚起来。

个人内部有一种极其有趣的直觉多样性,这种多样性源自于直觉明显敏感于在针对某些假定例子的描述中是否存在着**情感内容**。第二章中,我们已经看到了一个例子,回想一下肖恩·尼科尔斯和约书亚·诺布(2007 年)所发现的,人们对于因果决定论和道德责任的关系有着不同的直觉,这取决于他们正在评价的那个假定的例子有没有情感内容。当然,这种直觉敏感性并不局限于形而上学直觉,埃里克·乌尔曼(Eric Uhlmann)、大卫·皮萨罗(David Pizarro)、大卫·坦南鲍姆(David Tannenbaum)和彼得·蒂托(Peter Ditto)(2009 年)最近发现,一些道德直觉也表明它们受到了富有情感的内容的影响。① 在一项研究中,乌尔曼和他的同事要求人们去考虑朱迪斯·贾维斯·汤姆森(Judith Jarvis Thomson)(1985 年)的**胖子**例子的两个版本:

胖子

考虑一个我称之为**胖子**的例子,在这个例子中,你站在电车轨道上方的一座人行天桥上。你可以看到一辆失控的电车沿着轨道疾驰而来。你回过头,想去看一下这辆车冲向哪里,你发现人行天桥下面的轨道上有五名工人。怎么办呢?作为一名电车专家,你知道让一辆失控的电车停下

① 道德直觉似乎也受到了关于正在考虑那个假定的例子的人的事实的影响。赞姆佐和尼科尔斯(2009 年)最近发现道德直觉还表现出了性别对它们的影响。

> 来的一个切实的办法就是在它的轨道上扔一个非常重的东西。但是去哪里找这个东西呢？在天桥上，刚好有一个胖子，一个很胖的胖子站在你身边。他倚在栏杆上，看着电车。你要做的只是轻轻地推一把，他就会翻下栏杆，掉在电车的轨道上。你这样做是否被允许呢？

一些人被要求去考虑这一例子的一个版本，在这个版本中，一个行动者可以选择牺牲一个名为“奇普·艾尔沃斯三世”(Chip Ellsworth Ⅲ)的人来挽救100个哈莱姆[①]爵士乐队的人。其他人被要求去考虑另一个版本，在这另一个版本中，这个行动者可以选择牺牲一个叫“特伦·帕顿”(Tyrone Payton)[②]的人来挽救100个纽约交响乐队的人。引人注意的是，比起特伦· 79
帕顿，人们(尤其是那些自认为是政治自由主义者的人)更有可能认为牺牲奇普·艾尔沃斯三世是被允许的。[③] 针对这个不对称的结论，存在着一种解释，那就是那些自认为是政治自由主义者的人会倾向于显示出对采取如下这种态度的厌恶，这种态度可能会被认为是有偏见的(普兰特[Plant]和德维尼[Devine]，

---

① 哈莱姆，纽约市的一个区，位于曼哈顿东北的第96街以北，有大量黑人在此居住，20世纪20年代和30年代曾以其夜总会和爵士乐队闻名。这里说的应该是一个(也许是虚构的)黑人爵士乐团。——译者

② 特伦·帕顿通常应该是黑人的名字，而前面的奇普·艾尔沃斯三世通常应该是白人的名字，因此才有下面的种族偏见一说。——译者

③ 乌尔曼等人(2009年)用线性回归来测试场景的条件(特伦·帕顿 vs 纽约交响乐团=−1，奇普·艾尔沃斯三世 vs 哈莱姆爵士乐队=1)、政治倾向(1=非常自由主义，5=非常保守主义)和对效果论的支持程度(1=完全不支持，7=完全支持)的独立和交互影响。他们发现一种可靠的低阶效应：比起奇普·艾尔沃斯三世被牺牲的情况，他们在特伦·帕顿被牺牲的情况下更不愿意支持效果论($b=-0.19$，$SE=0.09$，$t[84]=2.24$，$p=0.03$)。对于那些自认为政治自由主义者的人来说，影响更为显著(低于平均数1 SD)：$b=-0.40$，$SE=0.12$，$t=3.27$，$p=0.002$。

1998 年;莫尼[Monin]和米勒[Miller],2001 年;诺顿[Norton]等人,2004 年)。这种甚至是对表面的偏见的厌恶,再加上这两个例子的不同的种族特性,导致人们(尤其是那些自认为是政治自由主义者的人)对这两个例子有着不同的情感反应,这些情感反应又导致了不同的道德直觉。①

我们已经看到,哲学直觉敏感于文化背景、性别和针对假定例子的描述中有无情感内容。最近,斯特西·斯温(Stacey Swain)、约书亚·亚历山大和乔纳森·温伯格(2008 年)发现了另一种直觉的敏感性:直觉还敏感于这个假定例子在其中被考虑的背景。具体说来,他们发现了实验主体有着不同的认知直觉,这取决于他们是不是要先去考虑其他的假定例子,考虑哪一个假定例子。比如,考虑一下凯特·莱勒(Keith Lehrer)的"特鲁普特"例子的如下版本:

特鲁特普

有一天,查尔斯被一块飞石击中了;结果他的脑子被"重新装配"了,他在估计他所处的地方的温度时总是对的。查尔斯并不知道他的大脑已经被改成这副样子了。几个星期以后,这个重新装配的大脑使他相信他房间的温度是华氏 71 度。除他的估计外,他没有什么认为现在温度是 71 度的理由。而事实上现在温度正是 71 度。

一些人被要求在评价其他任何假定例子前先去考虑特鲁特

① 那些自认为是政治保守主义者的人的直觉也无法做到不受影响。在一个独立的研究中,乌尔曼和他的同事发现,比起伊拉克反抗者做出一些间接涉及平民伤亡的行为来制止美国军队,那些自认为是政治保守主义者的人更有可能认为美国军队做出一些间接涉及平民伤亡的行为来制止伊拉克反抗者在道德上是可取的。

普例子。其他人被要求在考虑完一个关于知识的明确例子后去考虑**特鲁特普**例子：

> 凯伦是一名杰出的化学教授。今天早上，她在一篇发表在主流科学期刊上的文章中读到，将两种常见的地板消毒剂——Cleano Plus 和 Washaway 混在一起，就会产生一种对人体有害的毒气。实际上，这篇文章是正确的：将这 80
> 两种产品混在一起确实会产生出一种有毒的气体。中午的时候，凯伦看到一个门房将 Cleano Plus 和 Washaway 混在了一起，她朝他喊道："走开！将这两种东西混在一起会产生有毒的气体！"

还有一些人被要求先去评价一个明显不是知识的例子，然后再去评价**特鲁特普**例子：

> 戴夫喜欢用掷硬币来玩游戏。他有时候会有一种"特殊的感觉"：下一次投掷将会是正面朝上。当他有这种"特殊感觉"时，他有一半是对的，有一半是错的。在下一次投掷之前，戴夫有了那种"特殊的感觉"，这个感觉使他相信这枚硬币将会正面朝上着地。他投出了硬币，而且它落地的时候确实正面朝上。

我们发现，比起那些被要求在评价其他任何例子之前就去评价**特鲁特普**例子的人，那些被要求先去评价一个知识的明确例子，然后再去评价**特鲁特普**例子的人更不愿意在**特鲁特普**例子中作出知识的归属，而那些被要求先去评价一个非知识的明确例子，然后再去评价**特鲁特普**例子的人更加愿意在**特鲁特普**

例子中作出知识的归属。① 这表明**特鲁特普**直觉受到了这个假定例子在其中被考虑的背景的影响。

心理学家已经意识到了直觉的这种敏感性。在一个早期的研究中,列维斯·彼得诺维奇(Lewis Petrinovich)和帕特里西亚·欧尼尔(Patricia O' Neill)发现我们的一些道德直觉也敏感于次序和背景。他们要求人们去考虑**电车**例子的三个版本。②第一个版本如下:

开关

> 一列电车正在轨道上飞驰。电车前面的轨道上有五个无辜的人,如果电车继续往前开,他们将会被撞死。那条轨道的岔路上有一个无辜的人。电车的刹车坏了,然而人们可以打开
> 81 一个开关,让电车开到岔路上去。你是一个无辜的路人(不是铁路的职工,等等)。你可以打开这个开关,挽救五个人的性命,但是这会导致岔路上那个无辜者的死亡。你会怎么做?

第二、第三个版本只在无辜者的位置以及为挽救这五个人

---

① 更具体地说来,实验对象被要求对他们赞成或不赞成查尔斯知道房间里是华氏 71 度的程度给出评分(用一个 5 分制的李克特量表法,1 分=强烈反对,5 分=强烈赞成)。那些被要求在评价任何其他例子之前去评价**特鲁特普**例子的实验对象给出的平均分是 2.8 分,那些被要求先去评价一个明显的知识的例子然后再去评价**特鲁特普**例子的实验对象给出的平均分是 2.4;而那些被要求先去评价一个明显非知识的例子然后再去评价**特鲁特普**例子的实验对象给出的平均分是 3.2 分。这一差别在统计学上达到 p 值小于 0.05 的显著水平。赖特(Wright)(2010 年)和温伯格等人证实、复制并扩展了这些结论。赖特给出了一个关于直觉受到例子给出的顺序的影响的有趣解释。温伯格等人考察了人们的**认知需求**(need-for-cognition)(卡西欧普[Cacioppo]和佩蒂[Petty],1982 年)与直觉敏感于给出例子的顺序的关系。

② 对这一工作的有趣讨论,以及对这种敏感性在传统哲学方法中引发的方法论问题的讨论,可见辛诺特·阿姆斯特朗(2008 年)。

而需要做出的行为上有些差别。在第二个版本中——让我们称之为**按钮**例子——这个无辜者站在轨道上方的一座桥上，而为了挽救轨道工要做出的行为是按一个按钮，这样会在电车下面产生一个斜坡，使得电车跳出轨道，跃到桥上，撞死那个无辜者，然而会挽救那五个工人的生命。在第三个例子中——让我们称之为**推**的例子——这个无辜者和你一起站在桥上，为了挽救轨道工要做出的行为是将这个无辜者推到电车的轨道上去。彼得诺维奇和欧尼尔要求一些人按照前面给出的顺序（**开关——按钮——推**）来考虑这些例子，要求另一些人用相反的顺序（**推——按钮——开关**）来考虑这些例子。他们发现，在**开关**和**推**的例子中，比起出现在序列的最后面，当这些例子各自出现在序列的最前面时，人们更希望做出行动来挽救五个工人；①在**按钮**例子中，比起紧接着**推**的例子，当**按钮**例子紧跟着**开关**例子时，人们更希望做出行动来挽救五个工人。② 这表明了我们关于这些例子的直觉，就像我们关于特鲁特普案例的直觉一样，是敏感于这些例子在其中被考虑的背景的。③

---

① 实验对象被要求用一个从+5到−5的范围指示出他们是否愿意为挽救那五个工人的生命而行动。当开关的版本首先出现的时候，平均分是3.1分，当开关的版本排第三出现的时候，平均分是1.0分。胖子版本首先出现的时候，平均分是−0.86，胖子版本排第三出现的时候，平均分是−1.7。这一差别在统计学上达到p值小于0.005的显著水平。

② 和以前一样，实验对象被要求用一个从+5到−5的范围指示出他们是否愿意为挽救那五个工人的生命而行动。当按钮版本紧随着开关版本的时候，平均分是2.7分；当按钮版本紧随着胖子版本的时候，平均分是0.65分。这一差别在统计学上达到p值小于0.02的显著水平。

③ 彼得诺维奇和欧尼尔的研究在某种程度上受到了如下事实的限制，即他们问的是人们**将会做什么**，而不是在人们看来什么是**道德上允许的**或**道德上正确的**。正如赞姆佐和尼科尔斯（2009年）正确地指出的："关于你在某个情景中会做什么的判断可以与你的道德判断相分离。"近来关于道德直觉（尤其是关于**电车**例子的直觉）受语境影响的研究已经解决了这一问题。可见里奥等人（即将出版）和史怀泽格贝尔和库什曼（2011年）。

结果发现，哲学直觉敏感于一大堆我们未曾预料它们会敏感于的东西（敏感于是谁在考虑这些相应的假定例子，以及有没有某些种类的内容，以及这个假定例子在其中被考虑的语境），这一敏感性产生了**人与人之间的和个人内部的**直觉多样性。正如我们将会看到的那样，直觉的敏感性，尤其是那些未被预料到的和不受欢迎的直觉敏感性，为我们基于直觉的实践带来了另一种质疑，这些实践未来的成功将取决于我们更好地理解与接纳这些形式的直觉敏感性的能力。

4. 解释“限制主义”

82 这些研究并没有给出最终结论，但是它们确实有力地揭示了一些哲学直觉敏感于文化背景、性别、情感作用和语境。敏感性本身并不是问题。我们希望我们的证据是敏感的，但是我们希望它敏感于正确的东西。具体说来，我们希望我们的证据仅仅敏感于那些我们认为与一些主张的真假有关的东西，而这些证据应该为这些主张提供证据。使得这种直觉的敏感性**成问题**的是这种影响是**不受欢迎的**：似乎一些哲学直觉敏感于一些东西，而大多数哲学家并不希望它们敏感于这些东西，或至少从未希望过它们敏感于这些东西。大多数哲学家并不认为关于知识或意义的主张的真假取决于如下事实，这些事实与是谁在评价这些相关的假定例子有关。大多数哲学家并不认为关于道德责任或道德上正确的行为的主张的真假取决于我们的情感状态是否在发挥作用。同时，大多数哲学家并不认为关于知识或道德责任的主张的真假取决于那些主张在其中被评价的语境。当然，传统看来这些东西并不被认为是与哲学主张的真假有关的这一事实并不意味着它

们实际上是无关的。① 对于什么东西是与哲学主张的真假有关的，哲学家可能搞错了。但是，它确实给我们的基于直觉的实践带来了压力。至少哲学家面临着一个两难：要么我们必须解释为什么这种直觉的敏感性是受欢迎的，要么我们必须不再把这些哲学直觉当作证据来诉诸，并且对我们的基于直觉的实践增加一些**局部的限制**。（“限制主义”这个名称来自这一“两难”的“第二难”。）

无论我们选择哪一个选项，对我们的基于直觉的实践的这一质疑并不会在这里终止。这种直觉的敏感性不仅是不受欢迎的，它还是**预料之外的**。即使是现在，我们还缺乏预测这种直觉的敏感性在何时以及何处会出现的办法，同时，我们还发现我们自己站在一个不牢靠的认知立场上，我们怀疑**一些**直觉证据是成问题地敏感的，却无法可靠地预测**什么样的**直觉证据是成问题地敏感的。 83

乔纳森·温伯格（2007 年）将这种认知处境称为是“没有希望的”，但并不**完全**是这样的。我们缺乏发觉和纠正成问题的直

---

① 实际上，现在至少还存在着两种哲学理论，根据这些理论，这些东西（尤其是文化背景和语境）中的一些是与此相关的，它们就是相对主义和语境主义。根据相对主义的看法，某些哲学主张的真假取决于（至少部分地取决于）关于评价这个主张的人的事实。根据语境主义的看法，某些哲学主张的真假取决于（至少部分地取决于）这个主张在其中被给出和评价的语境。（近来对相对主义的讨论，可见博加桑和海勒斯，2006 年，近来对语境主义的讨论，可见德娄斯，1995 年；刘易斯，1996 年和科恩，1998 年）。虽然在某种程度上要评价这两个理论如何能够成功地解释为什么诸如文化背景、性别、情感内容或语境能够与哲学主张的真假有关超出了本文的范围，但是有两点是值得考虑的。首先，这些研究所揭示的东西是一些哲学直觉**受到了文化背景、性别、情感和语境的影响**。然而，相对主义者和语境主义者所需要的，是关于哲学直觉应该受到这些东西影响的某些独立的论证。换言之，他们需要某种从描述性主张过渡到规范性主张的方法。其次，不清楚的是，相对主义或语境主义（不论是各自还是放在一起）是否能解释为什么**所有这些东西**都是相关的。

觉敏感性的能力，但是这种状况并非是永久的。根据温伯格的看法，有四种东西将带来希望：外在的进一步的证据（证据来源**之间**的一致性）；内在的一致性（证据来源**内部**的一致性）；对边缘情况的探测能力（意识到某个来源的限制）；理论性的启发（当这些来源在运作的时候，意识到它们是如何运作的，当它们不运作的时候，意识到为什么它们不运作）。如果直觉并不是完全没有希望的，那么问题就变成了：为了恢复希望，我们该做什么。

恩斯特·索萨（2007 年 a）曾经提出，也许我们只是应该更加仔细而已。索萨比较了直觉证据和知觉证据，他写道：

> 当然，激发效应和框架效应（effects of priming, framing），①以及其他那些类似的情境因素会一般地影响直觉的认知状况，正如它们也会一般地影响知觉观察的认知状况。人们会认为，面对着针对知觉判断的那种影响时，维持知觉在认知上的重要性的方法也以一种类似方式适用于面对着针对直觉判断的那种影响时维持直觉在认知上的重要性的方法。结论是，我们应该**仔细**对待我们使用直觉的方式，而不是说那种直觉是无用的。（第 105 页）

这里的看法是：我们应该以处理关于知觉的认知地位的担忧的同一种方式来处理关于哲学直觉的认知地位的担忧。我们知道知觉也是成问题地敏感的（比如，考虑一下知觉错觉或幻觉），然而我们并不认为我们基于知觉的实践受到了这一事实的严重质疑。相反，对于我们使用的知觉证据是什么，以及什么时

---

① 关于激发效应和框架效应，即 priming effects 和 framing effects，国内学界已经有了很多介绍。——译者

候使用它，我们是很仔细的。

仔细是有益的，但是只有当我们知道仔细是什么意思的时候，那才是有益的，这就是直觉证据和知觉证据的类比失效的地方。当感官知觉搞错了的时候，我们对其有很好的了解，这反映在我们的知觉实践中，并且得到了一种针对那些掌管着我们的 84
知觉判断的机制的公共科学认识的增强。这就阻止了关于知觉那种成问题的敏感性的担忧演变成关于知觉证据的认知地位的全面担忧。然而，问题就在于，就直觉证据而言，我们并不处在一个相同的位置上。我们无法预测或解释直觉那成问题的敏感性。就直觉而言，这一点把我们放在了这样一个认知位置上，这个位置不同于我们相关于知觉的认知位置。在这种意义上，我们还不知道仔细是什么意思。学会仔细意味着发展出一种对直觉运作之方式的更好的认识。如果我们想要知道**什么样的**直觉证据是可用的，以及**什么时候**能够被使用，那么我们基于直觉的实践必须要去了解那些相应的心理学、认知科学以及受经验指导的心灵哲学。关于哲学直觉来自哪里，由什么样的机制产生，以及哪些因素在影响着它们，我们需要知道得更多。只有到那个时候，我们才有可能在基于直觉的实践中做得更为仔细。

5. 对限制主义的辩护：前车之鉴

这不是我们的基于直觉的实践第一次受到质疑，我们有必要问这样一个问题，即这种限制主义的质疑从过去的那些质疑中学到了什么。我们从过去那些针对我们的基于直觉的实践的质疑中学到的一个东西，就是一个成功的质疑不能依赖于一种过于狭隘的哲学观念。希拉里·科恩布利思(1998年，2002年，

2007年)有一个著名的观点,就是对我们的哲学实践来说,哲学直觉确实并不是十分重要的,或者至少不像很多哲学家所建议的那般重要。这里的看法是,由于我们的哲学直觉所提供的仅仅是关于我们的个人的或共有的概念的证据,又由于哲学家不应该对这样的东西过分感兴趣(或者至少他们的兴趣不应该仅仅局限于这样的东西),因此在哲学实践中哲学直觉并不扮演一个重要的角色。这并不是说它们在哲学实践中**不**扮演任何角
85 色。科恩布利思承认它们可以作为一种出发点,指出我们早期的哲学研究的方向;但是,这些出发点应该"为对于外部现象的更为直接的经验研究让路"(科恩布利思,1998年,第135页)。针对基于直觉的哲学实践的这种质疑的问题就在于它依赖于一种过分狭隘的哲学观念。当然,一些哲学家并不对我们关于事物的概念感兴趣,而对事物自身感兴趣,而且科恩布利思的"在对非心理的世界的哲学讨论中,哲学直觉并没有什么重要的角色要扮演(或者至少不能起到像其中的一个角色那般重要的作用)"的说法也许是对的。但是,其他哲学家对我们关于事物的概念感兴趣,而且这样做是对的。这一兴趣也许还伴随着一种对于事物自身的兴趣,同时还带有这样一种看法,即只有先理解了我们关于事物的概念,我们才能开始去理解事物自身。这种兴趣可能完全不同于对于非心理的世界的兴趣,它的着眼点是我们的哲学概念或那些掌管着我们的概念判断的心理机制。无论是哪种,即使最后发现对于那些不关心我们关于事物的概念的哲学课题来说我们的哲学直觉并不十分重要,要证明它们完全没有哲学上的重要性,那还需要更多的东西。

我们从过去对我们基于直觉的实践的质疑中学到的另一个东西,就是一个成功的质疑不应该过分强硬,以至于有破坏我们的**全部**证据实践的危险。让我们来看一下如下两个例子。第一

个例子涉及哲学直觉的可错性，可以被相当迅速地处理完。迈克尔·德维特（Michael Devitt）（1994 年）和凯瑟琳·埃尔金（Catherine Elgin）（1996 年）都提出哲学直觉不能被当作证据，因为它们是可错的。这个看法的问题应该是清楚的：所有一般认定的证据来源都是可错的，因此这一看法将会导致**全面**怀疑主义（global skepticism）。由于我们中的大多数人都认为全面怀疑主义应该被避免，因此我们不能仅仅因为它们是可错的就抛弃直觉证据，或任何一种证据。第二个例子涉及哲学直觉的可信性，更精确地说来，是我们如何才能**相信**它们是可靠的。罗伯特·卡明斯（Robert Cummins）（1998 年）提出哲学直觉不能被视为证据，因为我们并没有什么独立的正当性证明来将它们视为通向真理的可靠向导。除非依赖哲学直觉，否则我们无法建立哲学直觉的可靠性。把我们关于认知内在主义的所有可能 86
的担忧都放在一边，这一看法的问题在于它也会通向全面怀疑主义（普斯特，2000 年，戈德曼，2007 年）。为了搞清原因，让我们先从知觉开始吧。建立知觉的可靠性的最自然的方法，就是以过去的可靠知觉为依据给出某种归纳论证，但是这种归纳论证有赖于用一些知觉来衡量另一些知觉。实际上，我们是用知觉来建立它自身的可靠性的。对于内省来说，情况也是一样的，而且对于我们证据的任何一种基本来源来说，或许都是这样的。我们不能建立我们证据的任何一种基本来源的可靠性，除非依赖其他的证据来源，因此如果某种认知的环环相扣是不被允许的，那么有危险的不仅仅是我们的基于直觉的实践。

从关于认知的环环相扣的担忧中学到的另一个东西，就是对于我们的基于直觉的实践的质疑不应该要求那种在认知上不可能的东西。比如，让我们再考虑一下这样一个看法，即哲学直觉不能被视为证据，除非对于它们是通向真理的可靠向导这一

点我们有着独立的正当性证明。正如前面已经看到的，建立某种假定的证据来源的可靠性的最自然的方法，就是用过去的准确报道（这报道是由那种证据来源给出的）为依据给出某种归纳论证。除非我们承认某种认知的环环相扣，否则这就需要我们用其他假定的证据来源的报道来检查那种假定的证据来源的报道。因此，如果我们要以一种避免认知的环环相扣的方式来建立哲学直觉的可靠性，那么我们必须诉诸于某种假定的证据来源（我们可以称这一假定的证据来源为 E_1），而不是哲学直觉。当然，除非我们建立了它的可靠性，否则我们无法将某物视为证据，这一限定也适用于 E_1，正如它同样也适用于直觉。这样一来，如果为了建立哲学直觉的可靠性我们诉诸于 E_1，我们必须首先以其他假定的证据来源（我们可以称这一假定的证据来源为 E_2）为基础来建立 E_1 的可靠性。我们很容易看清楚这里出现的问题。我们被指派了一个不可能完成的任务。我们不能将哲学直觉视为证据，除非我们建立了它们的可靠性；但是，为了
87 建立**任何**一种假定的证据来源的可靠性，我们不得不先建立起其他**一些**假定的证据来源的可靠性。这是不可能完成的。

我们认识到了针对我们基于直觉的实践的更早的质疑所面临的一些问题。关于一个成功的质疑必须是什么样子的，这些问题为我们提供了一些准则：成功的质疑不应该依赖过于狭隘的哲学观念；它不应该太强硬，以至于会导致全面怀疑主义；它不应该要求那些在认知上不可能的东西。那么，实验的限制主义对这些东西学得怎么样了呢？似乎学得很好。尽管实验的限制主义似乎向一个特别的哲学观念**推进**了一步——这个哲学观念认为我们的哲学直觉告诉我们的仅仅是关于我们个人的或共有的概念的东西，但是实验的限制主义并未被任何特定的哲学观念所**推动**。还有，实验的限制主义似乎既没有威胁到我们其

他的证据实践的地位并以此引发全面怀疑主义,也没有要求那些在认知上不可能的东西。我们大多数的证据实践是有希望的,因此并没有面临我们的基于直觉的实践近来面临的那种方法论质疑。就我们的大多数证据实践而言,我们对应该使用什么样的证据以及何时使用这种证据有着很好的了解,因此我们的基于直觉的实践所面临的这种方法论质疑完全局限于那些基于直觉的实践。而且,如下这一点是完全可能的,即限制主义者认同的那种方法论规定在认知上并不是非理性的:我们只是被要求把我们对于哲学、哲学证据和我们的基于直觉的实践的看法与某些种类的直觉敏感性的存在调和起来;同时,我们还被要求花费更多的时间和精力来仔细思考直觉证据的性质,它来自哪里,什么样的机制掌管着它的产生,是什么因素在影响着它。

6. 结　论

我们认识了问题,认识了这个问题给我们的基于直觉的实践带来的挑战。一些直觉证据敏感于一些东西,我们既没有预料到这些东西,也许也不希望它们是这样的,作为回应,我们必须要么对直觉的这种敏感性表示欢迎并借此将这些问题解释 88
掉,要么对应该使用哪种直觉证据以及什么时候使用它附加一些限制。无论是哪种,似乎都需要我们花费更多的时间。我们不但要花时间努力思考我们的哲学实践,还要思考我们自身和我们的哲学直觉的本性。而且,这意味着无论哪一种选择都需要更多的实验哲学。

第五章
为实验哲学辩护

1. 引　　言

要求变革总会受到抵制，尤其会受到这样一些人的抵制，这 89
些人已经在正统哲学上投入了很多东西。这样一来，实验哲学并没有被普遍接受这一点也许是不足为奇的。那些在思考哲学及哲学方法论的更为传统的方式上投入了很多东西的哲学家并不能泰然处之，反而试图发动针对传统的一系列辩护活动。在本章中，我们将考察哲学家在为传统辩护时采用的三种更为常见的策略。（还有其他一些为传统辩护的方式，包括基于方法论的依据来质疑实验哲学。我们将在后记中审查这种方法论质疑。）这些策略的目标是实验哲学家的工作，它们提出这些工作是**不切题的**——要么是因为实验哲学注重的是错误的哲学直觉的类型，要么是因为它过于注重哲学直觉了。

2. 谁的直觉是重要的？

也许不是所有的哲学直觉都是生来平等的。这个相当直白 90
的看法变得越来越有吸引力了，这部分源自于它似乎对实验哲学作出了一种同样直白的回应。这种回应是这样的。不管实验

哲学家在研究什么,他们都没有在研究那种与基于哲学直觉的实践相关的直觉证据,因此,这些研究的结论对那些实践说不出什么东西,或者贡献不出什么东西。

这一回应有一个特别流行的版本,这个版本作为**专业知识辩护**(expertise defense)①而为人所知。根据专业知识辩护②的看法,我们对哲学直觉的兴趣应该被理解为是对**哲学家**的直觉的兴趣。关于专业知识辩护,基尔克·路德维希(2007 年)在讨论哲学直觉在语言哲学中扮演的角色时给出了一个精彩而简单的阐述:

> 我们不应该事先期待那些未受训练的实验对象会良好地给出判断——这些判断是对那些包含专名的语义学难题的小场景的回应,因为这是一个相当复杂的领域,在这里,我们日常的词汇并不是十分准确的。我们应该期待的是,哲学语义学领域的相关专家将会更好地给出答案,这些答案关注的是这些例子的正确特征以及这些特征应该与之对

① 这个名称来自温伯格等人(2010 年),他们首次对这种辩护的优劣进行了详细的讨论。辛格(1972 年、1982 年),德维特(1994 年、2006 年、2010 年),威廉姆森(2005 年、2007 年),海勒斯(2006 年),路德维希(2007 年),古伦德曼(Grundmann)(2010 年)和霍瓦特(Horvath)(2010 年)提出了各种版本的专业知识辩护。有趣的是,它的主要原则(比起大众的直觉,哲学家应该更加关注专家的直觉)也可以在斯蒂克和尼斯比特(1980 年)对反思平衡的方法以及推理规则的确证的讨论中被找到。由于很多不被接受的推理规则与大众的推理实践处在一种反思平衡中,他们主张一个推理规则只有在与经过训练的权威或专家的推理实践处于一种反思平衡中时,才是被确证的。然而,重要的是要知道斯蒂克(1988 年)收回了对专家反思平衡的这一早期辩护。

② 此处的"辩护"是对传统哲学方法的辩护,同时也是对实验哲学的质疑,因为实验哲学背离了传统哲学方法,二者的对立使得对后者的辩护成为了对前者的质疑。——译者

> 应的东西……要提倡的是这样一门学科的进步，在这门学科中，进行思想实验的一般专业知识将会得到灌输，同时概念考察的不同领域的专业知识将会发展和完善。确实有这样一门学科。它的名字是哲学。由于他们的训练和专业知识，哲学家最适于在他们的专业知识领域中做思想实验，并且解决这样一些方法论的和概念的难题，这些难题是在试图搞清楚概念（我们就是带着这些概念面对这个世界的）的复杂结构时产生的。（第 150—151 页）

斯蒂夫·海勒斯(Steven Hales)(2006 年)也强调教育和经验，他写道：

> 直觉敏感于，而且应该敏感于相关领域的教育和训练。 91
> 比如，职业科学家的物理学直觉要比大学生或那些公交车站的随机人员的直觉可靠得多。科学家有物理学直觉，并且依赖这种直觉，这些直觉是经过训练的，是熟练的，是有根有据的，因为这些理由，它们是真理的优秀指针。以这种同样的方式，职业哲学家的模态直觉①要比缺乏经验的学生或“大众”的直觉可靠得多。（第 171 页）

这里的共同看法似乎是这样的。比起大众的哲学直觉，我们应该对专家的哲学直觉更感兴趣。毕竟，哲学家有着更好的概念和理论，或至少对于相关概念有着更好的理解，他们对这些概念已经思考了很久，也思考得很努力，而且他们接受过训练，

---

① 原文为“modal intuition”，这里翻译成模态直觉，“模态”的意义见“模态状况”，第 29 页注③。——译者

知道如何最好地阅读和思考哲学的思想实验，这些实验要求我们去使用这些概念和理论。当然，这一点使专业的哲学直觉比大众的哲学直觉具有更多的理论价值。

虽然这是一个很有吸引力的看法，但是最后发现要确定“谁”关于“什么”以及“什么时候”具有专业知识，是很困难的。似乎只有某些种类的训练才能有助于改善任务的执行，而且即使这样，也只是改善某些种类的任务，我们有理由担忧哲学训练可能不是那种正确的训练，哲学的思想实验可能也不是那种正确的任务（比如，可见桑蒂[Shanteau]，1992 年。埃里克森[Ericsson]等人，2006 年，以及温伯格等人的讨论，2010 年）。因此，我们不能简单地假设专业的哲学直觉比大众的哲学直觉具有更多的理论价值。相反，我们必须仔细审查认为它们应该具有更多理论价值的可能理由。

我们认为专业哲学直觉比大众哲学直觉具有更多的理论价值的一个可能理由是哲学家有着更好的概念和理论，或者至少对相关的概念和理论有着更好的理解。彼得·辛格（1972 年）在他对道德的专业知识的讨论时提出了这一看法，他写道：

92 > 比起普通人，似乎道德哲学家确实有着重要的优势……他在道德哲学中的特有经验让他理解了道德概念和道德论证的逻辑。某人在缺乏对相关概念的准确理解的情况下参与了道德论证，产生了极大的困惑——在最近的道德哲学中，这一可能性得到了充分的强调，这里已无须说明了。明晰性本身并不是目的，而是良好论证的助手，而对明晰性的要求是道德哲学家已经认识到的某种东西。（第 117 页）

希拉里·科恩布利思(2007 年)也强调拥有恰当的理论认识的好处,他写道:

> 比起那些没有受理论指导的判断,科学中那些受理论指导的判断也许会告诉我们更多的东西,因为准确的背景理论导致更为准确的受理论指导的判断。不受理论指导的观察者和熟练的科学家都在试图把握独立存在的现象,在这个任务中,准确的背景理论是有帮助的。比起门外汉,专家是更好的观察者。如果哲学理论的构建情况是类似的,那么我们应该认为哲学家试图刻画的东西并不是他们的概念,更不是大众的概念,而是某些精神之外的现象,比如知识、确证、善、公正,等等。比起大众的直觉,职业哲学家的直觉更适合于把握那些现象,因为关于这些现象和它们的概念,他们已经思考了很久,也思考得很努力,同时,如果一切如应当发生的那样正常发挥作用,那么比起那些新手,职业哲学家更接近于对那些现象的准确刻画。(第 35 页)

让我们先从这一看法开始,即哲学的专业知识由具有更好的概念或至少有着更好的概念性理解构成。有几种阐释这一看法的办法。一种办法是提出哲学讨论一般要涉及**专业概念**,由于哲学家的哲学训练,他们对这些概念享有特权。这种说法的问题在于,要说大多数哲学讨论都涉及专业概念,这似乎并不是真的。[①] 当然有一些哲学讨论是这样的;那些关于**有效性**(validity)或**有保证的可断定性**(warranted assertability)便浮现 93

① 杰克曼(Jackman)(2001 年),尼科尔斯(2004 年 a),戈德曼(2007 年)和诺布及尼科尔斯(2008 年)提出了类似的观点。

在了脑海里。但是很多哲学讨论涉及的却是日常概念，而且这样做是有很好的理由的。对这些日常概念的担忧恰恰是最先引发这些哲学讨论的东西。如果这些讨论用纯粹专业词汇来表达，那么它们将会失去与引发这些讨论的日常概念的联系。比如，我们很难想象，关于那种涉及我们的日常概念（正是这些概念激发了很多当代的知识论）的认知担忧，关于知识的某些纯粹专业概念的哲学讨论能够告诉我们很多东西。

如果哲学的专业知识并不由具有更好的概念这一点构成，那么也许它由具有对日常概念的更好理解这一点构成。[①] 比如，也许哲学家有能力作出更加准确的概念区分。虽然这一点可能会让专家的哲学直觉比大众的哲学直觉更具有理论价值，然而重要的是要注意到，那些关于相互比较的概念能力的问题正好就是那种很适合于用实验哲学来帮助我们研究的问题。比如，约书亚·诺布和肖恩·尼科尔斯（2008年）写道：

> 专业知识的反驳论证的这一形式带来了一些有趣的问题，但是我们不知道它如何才能成为对实验哲学实践的反驳。相反，我们很希望知道哲学家与常人的更多的区别，对我们而言，要发现这些区别的最好方法就是去做一些实验……还有，即使我们发现了哲学家与常人之间的重要区别，也很难得出结论说来自大众的数据就是无关紧要的。应该说，数据的整个模式可能会告诉我们一些重要

---

① 路德维希（2007年）提出了这个建议，他是这样提到专业知识的："它和获得新概念没有关系！它是用涉及那些概念的问题来进行反思训练，从而获得对概念结构的更高的敏感度。"（第149页）

的东西，这些东西是与哲学问题的终极来源有关的。（第9页）

简言之，哲学家对日常概念有着更好的理解这一看法会导致**更多的**与我们的哲学直觉有关的实验工作，而不是更少。

它还会导致一个问题。哲学家对日常概念有着不同的理解，对于这一点的证据并不是对于他们对那些概念有着更好的理解的证据，因此，为什么我们会认为他们有更好的理解呢？一般说来，答案与哲学的教育有关——在某种程度上，哲学的教育 94
会改善我们的概念性理解。但是，这是如何发生的，这一点是不清楚的。乔纳森·温伯格、查德·贡纳曼、卡梅隆·布克纳和约书亚·亚历山大(2010年)指出，大多数哲学家认为那是借助某种试错(trial and error)的过程而发生的，在这里，学哲学的学生将他们的概念判断与某些约定的标准相比较，从而训练他们的概念能力。这里的一个假设是学哲学的学生用先前被证实的哲学直觉来训练他们的概念判断，但是这只会引来解释的回溯。**那些**哲学直觉又是如何被证实的呢？另一个假设是学哲学的学生用已经建立起来的哲学理论来训练他们的概念判断，但是这也会引起解释的回溯。如果学哲学的学生要用已经建立起来的哲学理论来训练他们的哲学直觉，而专家的哲学直觉在建立这些理论时扮演着一个重要的角色，那么我们仍然要疑惑于专家的哲学直觉是如何被证实的。当然，这里的问题在于，我们的哲学直觉不会收到任何类似于那种客观回馈的东西，为了确实地帮助我们改善我们的概念理解，这些东西似乎是必需的。实际上，事情会变得愈加糟糕。正如温伯格等人(2010年)所解释的：

> 有很多种心理机制可以干扰一般意义上的人类认知——以偏概全(overgeneralization)、过分自信、认知失调、归属错误(attribution error)、信念的偏见、固执的信念(belief perseverance)等等,这些机制很可能导致哲学家相信我们有一致的直觉,即使实际上在很多地方我们只是在系统地重新确认早期的印象,并且错误地将我们专业的成功(比如在辩论和出版物中的成功)归结于它们的有效性。(第 341 页)

我们不仅有理由担心哲学教育并不是很有帮助,还有理由担心我们并没有很好地被引导去认识这一点。

也许哲学的专业知识由对某些哲学理论和原理的掌握构成。希拉里·科恩布利思(2007 年)所想到的东西便与此类似,他写道:

> 95 不受任何理论指导的直觉,或者极少受到大众所熟知的理论指导的直觉,就像对科学中的相关背景理论完全无知的调查者给出的观察一样,[在哲学中]是没有用处的。一些人对相关的理论是如此无知,以至于他们的信念集不包含任何可能会影响那些观察的理论,在科学中,这些人的观察对我们没什么帮助……我们应该去把握那些前理论的直觉这一建议似乎把那些无知而幼稚的人的直觉放在了那些负责而有见地的调查者的直觉之前了。我不明白为什么这个看法在哲学中要比在科学中好一些。(第 34 页)

这个看法似乎是,哲学理论可以帮助塑造我们的直觉,其方

法可能是让一个给定的假定例子中的某些特征变得显著起来，或指导我们去解释那些特征。[①] 尽管这是一个具有吸引力的看法，但是它面临着几个难题。首先，或许经过哲学理论过滤的那些判断甚至不能被算作哲学**直觉**；至少根据某些看法（比如林奇，2006 年），为了让过滤之后的判断被视为真正的哲学直觉，这个过滤的过程应该是无意识的，或者是无法借助内省来把捉的。其次，信奉某种理论可以起到澄清的作用，但同样有可能起到搅浑的作用。专家的哲学直觉是受理论指导的，这一事实并不能确保它们比大众的哲学直觉具有**更多的**理论价值；实际上，正是由于那个理由，它们可能会具有更少的理论价值。科恩布利思认识到了这一可能性，但是他认为这仅仅意味着我们应该尽我们所能，来确保那些影响我们的哲学直觉的理论是准确的。尽管这对于基于直觉的哲学实践来说意味着什么还是不清楚的。如果我们对理论的信奉塑造了我们的哲学直觉，那么很难弄明白我们的哲学直觉如何能够**独立地**帮助我们去评价这些理论的准确性。对于像科恩布利思这样的人（他认为我们基于直觉的实践只有很有限的哲学价值）而言，这并不是一个问题。但是，有这样一些人，他们认为我们是以哲学理论解释我们的哲学直觉的能力为依据而提出哲学理论的，认为我们是以这些理论 96
与我们哲学直觉的总体一致性为依据来为它们的真辩护的，

---

① 在如下的例子中，路德维希（2007 年）似乎持有某些类似的想法：史密斯先生是一个二十多岁的正常的健康人，他现在正倚着起居室的窗户站着。听到房间另一头沙发边上靠墙的桌子上的电话响了。他走过他和电话之间的地板，来到桌子旁，把电话接了起来。如下哪一点是真的：（A）史密斯先生试图穿过起居室；（B）史密斯先生并不试图穿过起居室；（C）既不是（A）也不是（B）。按照路德维希的看法，我们**觉得**（A）是正确答案，因为这个答案最适合我们那个最好的故意行为理论，根据这一理论，一个人即便在那些他将会成功这一点几乎是确定的例子中，也能够试图做某事。

并且是以它们与我们哲学直觉的相符为依据来确证我们的哲学信念的。对于这些人来说，它确实给出了一个重要的质疑。

让我们稍微转换一下思路。也许哲学的专业知识仅仅在于如下事实，即哲学家花费了更多的时间来仔细思考哲学问题。安提·库皮尼(2007 年)采纳了某些类似的看法，他主张比起**表面的**哲学直觉，**坚实的**哲学直觉具有更多的理论价值。坚实的哲学直觉的一个标志，就是它们是在**足够理想的条件下**形成的，在这种条件下，我们不仅有足够的时间来仔细地检查和评价关于假定例子的判断，还可以检查和评价这些例子本身，以及我们对理论的信奉对我们在那些例子中发现的相关细节的可能影响。[①] 这里的潜在看法是，**反思性的**判断更有可能是正确的(比如，可见索萨，1991 年)，而且如果哲学家花费更多的时间来参与这种反思实践，那么认为比起大众的哲学直觉他们的哲学直觉具有更多的理论价值似乎就是自然而然的。

然而，正如希拉里·科恩布利思(2002 年，2010 年)和乔纳森·温伯格以及约书亚·亚历山大(即将出版)所指出的，反思和可靠性之间的关系并不是如此简单的。是的，确实存在反思帮助我们改善我们的判断的时候。有时候，反思可以帮助我们认识到我们对于某个具体例子的相关细节有了错误的认识或没有认识，认识到我们错过了某些相关的细节，或者我们最初关于哪些细节是相关的判断受到了我们对理论的信奉的干扰。但是，反思也很容易成为一个回音室，仅仅去认可我们最初给出的判断，不管这些判断是什么。它可能有助于增加我们对那些判

① 可见列文古德等人(2010 年)的一个有趣的经验研究，这一研究揭示了哲学家比大众更善于反思。

断的信心，而不会增加它们的可靠性。这里的一个问题是，我们的概念判断的形成所涉及的那些认知过程是无法借助内省通达的（尼斯比特和威尔森，1977 年；威尔森，2002 年）。① 另一个问题是，我们倾向于过分自信于我们自己的可靠性；我们高估了我们的信念以相关的细节为依据的程度，低估了我们关于哪些细节是相关的决定受到我们对理论的信奉的干扰的程度。然而，最大的问题也许是，当我们在实际上试图进行库皮尼心目中的那种反思时，一大堆认知偏见会妨碍到我们（比如，可见沃森 97
[Watson]，1960 年；尼克尔森[Nickerson]，1998 年；巴隆[Baron]，2000 年）。事实上，即使认识到这一点也不会有所裨益，因为甚至对于那些已知的认知偏见，我们在进行校正的时候也会表现得非常糟糕（普罗宁[Pronin]等人，2002 年）。简而言之，要说反思足以帮助支持专家的哲学直觉比大众的哲学直觉更有理论价值这一看法，这是完全不清楚的。②

还有另一个可能性，也许哲学的专业知识由某种**程序性**知识构成——某种从我们的哲学教育的过程中发展出来的特殊的

---

① 科恩布利思（2002 年）给出了两个很好的例子。第一个例子是尼斯比特和威尔森（1977 年）对位置效应（position effects）的研究。尼斯比特和威尔森发现人们对生活消费品的估价受到了那些要被评估的货物的相关位置的影响。即使被要求去思考一下是什么因素影响了他们的估价，人们也没有意识到相关位置对他们的估价的影响。第二个例子是特沃斯基和卡尼曼（1974 年）对锚定效应（anchoring effect）的研究。在这项研究中，实验对象被要求去估计非洲国家在联合国中的比例。当被问及这个比例是大于还是小于 10%时，人们的平均估计是 25%，当被问及这个比例是大于还是小于 65%时，平均估计是 45%。虽然很明显实验对象受到了最初的参照点（"锚"）的影响，然而，实验对象完全没有意识到这个参照点对他们的估计的影响。

② 实际上，温伯格等人（即将出版）让我们有理由担心即便是反思直觉也展现了前面章节中提到的那种成问题的直觉敏感性，尽管其方式可能略有不同。

"能力知识"(know-how)。① 如果哲学家在如何最好地阅读和思考哲学思想实验上受过更好的教育,那么这可能会自然而然地使他们的哲学直觉似乎比大众的哲学直觉更有理论价值。那么,这种程序性的专业知识看上去会是怎么样的呢?这里有两种可能性。恩斯特·索萨(2009 年)最近提出哲学思想实验一般采用的那些小场景需要读者输入某些信息,而这些信息并不明确包含在那些段落之中。以这样的方式阅读哲学思想实验就类似于阅读小说:②

> 当我们阅读小说的时候,我们输入了很多并不包含在文本中的东西。当我们在跟从作者的引导时,我们在自身的想像建构中输入了很多东西,它们通常是对物理和社会结构的建构所预设的东西。(第 107 页)

如果这是对的,那么我们也许可以用我们恰当地获得某个特定小场景的相关细节的能力,我们恰当地填补那些并未明确包含在小场景中的细节的能力来解释哲学的程序性专业知识。蒂莫西·威廉姆森(2007 年,2011 年)给出了另一种程序性专业知识。根据威廉姆森的解释,思想实验涉及带有反事实前提的有效演绎论证,而我们应该用一种想象模拟、背景信息和逻辑的混合体来评价这些反事实前提。威廉姆森(2011 年)试图将思想实验的任务分解为明确的子任务:我们必须阅读并消化对场
98 景的描述,判断这个被描述的场景中的情况是什么,判断这个场

① 有的学者翻译成"技能知识"。——译者

② 市川和贾维斯(2009 年)以大卫·刘易斯(1978 年)关于"小说中的真"(truth in fiction)的工作为基础,考察了哲学思想实验和小说的相似性。坎普(Camp)(2009 年)也有趣地讨论了想象在哲学思想实验中的作用。

景是否可能，并且确定这个前提是否能推演出结论。如果对思考思想实验的过程的这种重构是正确的，那么我们也许可以用我们在某个特定小场景中挑选出相关细节的能力，从事反事实的推导的能力和做出某些逻辑推论的能力来解释哲学的程序性专业知识。这两者似乎都是哲学专业知识的相当有趣的图画，但是重要的是要明白关于相互比较的程序性专业知识的问题，正如关于相互比较的概念能力的问题一样，恰好就是那种很适合于用实验哲学来帮助我们研究的问题。因此，很难把哲学家有着更多的程序性能力知识这样一种可能性转变成消除实验哲学的哲学重要性的理由。

我们已经审查了几个认为专家的哲学直觉比大众的哲学直觉具有更多的理论价值的可能理由：哲学家有着更好的概念和理论，至少对相关概念和理论有着更好的理解，哲学家对这些概念和理论已经思考了很久，也思考得很努力，哲学家在如何最好地阅读和思考哲学思想实验（这些实验让我们去运用这些概念和理论）这一点上接受过训练。至少现在看来，这些理由没有一个是特别具有说服力的。至少我们需要做更多的工作来解释为什么我们应该认为专家的哲学直觉比大众的哲学直觉具有更多的理论价值。这种工作的大部分应当是经验性的，并且要仔细地研究大众的哲学直觉、专家的哲学直觉，以及与思考哲学思想实验和产生哲学直觉相关的那些认知机制。①

---

① 有趣的是，对于专业知识辩护的拥护者而言，那些关于专家的哲学直觉的工作看上去并不很有说服力。史怀泽格贝尔和库什曼（2011 年）和舒尔兹（即将出版）等人的工作似乎揭示了专家的哲学直觉也展现了大众的哲学直觉所展现的那种成问题的直觉敏感性。

3. 什么样的直觉是重要的？

当然还有其他一些削弱实验哲学的重要性的办法。一个越来越流行的方法就是把我们的注意力从**谁的**直觉是重要的转移到**什么样的**直觉是重要的。乔纳森·温伯格和约书亚·亚历山
99 大(即将出版)将此称为“**厚辩护**”(thickness defense)，因为它一般要采用本书第一章中简述的那种哲学直觉的“厚概念”。厚辩护的基本策略就是提出不管哲学直觉是什么，它们都不是那种可以用实验研究来考察的东西——对于实验哲学家而言，那些实验研究已经是他们惯用的东西了。基尔克·路德维希(2007年)给出了厚辩护的一个相当清楚的例子：①

> 要给出的第一个观点是……人们对有关思想试验中采用的那些场景的调查的回应**本身**并不是直觉，它们**本身**并不是这样的判断——这种判断惟一地表达了实验对象在对场景作出回应时使用那些与场景相关的概念的能力……我们知道这些回应不是(至少不全是)直觉，面对这些回应，要做的事情就是试图将能力的贡献从其他因素中分解出来。为了能够自信地说出是什么因素在起作用，需要我们了解那些可能会对回应产生影响的各种因素是什么，还需要关于每一个实验对象的足够信息。很明显，在进行这些调查的情况下，我们没有这类信息。(第144—145页)

---

① 库皮尼(2007年)、库伦(2010年)、本森(即将出版)也提出了厚辩护的一些版本。

这里的基本看法是这样的，因为不是所有的调查回应都揭示了我们的哲学直觉，实验哲学家承担上了这样一个责任，那就是要区分哲学直觉和其他种类的精神状态（那些回应也可能揭示了这些精神状态），目前，实验哲学家的方法使他们无力承担这一责任。

这是一个颇有吸引力的策略，特别是把前面章节中提到的那种方法论质疑也加进来以后。如果实验哲学家没有在研究那种正确的东西，那么无论他们引起了何种方法论的担忧，这些担忧都不是关于哲学实践中**实际**采用的方法的担忧。然而，正如温伯格和亚历山大（即将出版）所指出的，只有当某些条件被满足时，这一策略才是有效的。这些条件一个与哲学直觉追踪真理的倾向有关，一个与我们在实践中成功地确定在我们的精神状态中有哪些精神状态被算作真正的哲学直觉的能力有关。①若最后发现，哲学直觉并不比实验哲学家在研究的其他那些精神状态更有可能追踪哲学真理，那么，发现实验哲学家从来就没 100
有在研究哲学直觉就不会让人有一点点安慰。类似地，除非我们有所需的办法在实际的哲学实践中挑选出那些正确的精神状态，否则发现实验哲学家没有在研究那种正确的精神状态也不

---

① 温伯格和亚历山大（即将出版）将这两种条件分别称为**写实主义条件**（veritist conditions）和**方法论条件**（methodological conditions）。他们主张存在两种重要的写实主义条件：一种是免疫条件（immunity condition）——实验哲学家研究的那些精神状态所展示的证据多样性和成问题的证据敏感性不应该（在任何一种令人担忧的程度上）影响到哲学直觉；另一种是**希波克拉底条件**（Hippocraticity condition）——其他种类的成问题的证据敏感性也不应该（在任何一种令人担忧的程度上）影响到哲学直觉。他们还提出存在两种重要的方法论条件：一种是**可明确表达条件**（manifestability）——我们可以说出哪一种精神状态是真正的哲学直觉；另一种是**当前实践条件**（current practice）——我们可以用通常与传统分析哲学相关联的方法来做到这一点。

会让人有一点点安慰。问题在于没有一个版本的厚辩护**明确**符合这两个条件。一些版本的问题在于它们确实没有给出什么东西，它们以这样一种方式来处理哲学直觉，以至于我们要么有充分的理由来担心哲学直觉不会追踪哲学真理，要么缺乏那种将它们和其他种类的精神状态区分开来的办法。但是，最一般的问题在于厚辩护的大多数版本没有回答如下问题：这些条件有没有得到满足。而且，这意味着我们需要更多的实验工作，而不是更少，即使那种实验工作需要更加仔细地确定被研究的是那些正确的精神状态。

4. 直觉重要吗？

也许问题并不是**谁的**直觉是重要的，或甚至是**什么样的**直觉是重要的，而是直觉**到底**重不重要。实验哲学家认同这样一个观点，即哲学直觉在当代哲学中扮演着一个重要的角色，它提供了要被我们的哲学理论所解释的材料，提供这些理论为真的证据，提供相信它们为真的理由。即便对于这样一些实验哲学家而言——这些哲学家认为，关于我们的心灵运作的方式和思考哲学问题的方式，我们知道得更多一些，就会引发关于直觉在哲学实践中扮演的角色的担忧——情况也是这样的。简而言之，实验哲学家严肃地对待哲学直觉，这意味着实验哲学的哲学重要性取决于，至少部分地取决于我们的哲学直觉的重要性。如果最后发现我们的直觉没有哲学上的重要性，那么实验哲学将不得不走到这样一个不幸的位置上
101 去——它严肃地对待这样一种思考哲学的方式，这种方式从一开始就不值得被严肃地考虑——实验哲学将会是没有哲学价值的。

顺着这些思路，主张哲学直觉在当代哲学中没有扮演什么重要角色的说法就变得流行了起来。马克斯·杜尔歇(Max Deutsch)(2010年)提倡这一看法，提出虽然**反例**(counterexamples)在当代哲学中扮演着一个重要的角色，但是我们关于这些反例的直觉并没有扮演这种角色。杜尔歇认为，某人是否拥有一个给定的反例为真的直觉，这纯然是一个心理学问题，忘记这一事实才能让实验哲学家在讨论哲学方法论时有所收获。乔纳森·市川(即将出版)也反对当代哲学依赖于我们的哲学直觉这一看法，并提出虽然哲学家依赖**直觉命题**，但是认为这意味着他们依赖直觉作为证据甚或他们依赖直觉命题是因为它们是直觉性的，这就是错误的了。按照市川的看法，当哲学家诉诸某些命题的直觉性质时，他们试图用它来论辩，而不是将它们作为证据。

哲学直觉是不是**真的**在当代哲学中扮演着一个重要的角色，这是一个社会学问题，而且在很大程度上既依赖于我们如何解释我们对"**我们会说点什么**"或"**事情对我们而言似乎是怎么样的**"[①]的诉诸，还依赖于在我们看来当哲学家诉诸那种东西的时候他们要做什么。我们很快就要来讨论这第二个问题了，但是在开始之前，有必要花一点时间来谈一谈哲学家谈论直觉的方式。比如，杜尔歇(2010年)提出，哲学直觉并没有在哲学中扮演着一个重要角色，因为大多数的哲学讨论并不包含对我们的哲学直觉的明确援引，即使是在如下这类讨论中——这些讨论通常被认为是关于哲学直觉在哲学实践中应该扮演核心角色的典型例子——情况也是

---

① 这是典型的针对小场景的直觉回应。——译者

一样的。[1] 然而,基于阿尔文·戈德曼(2007 年)所阐述的如下这些理由,这一步似乎走得太快了:

> 作为一个历史事实,哲学家并不总是用直觉的语言来描述他们的方法论。实际上,这似乎是近期才出现的做法。雅克·辛提卡(Jaakko Hintikka)(1999 年)将对“直觉”的哲学用法追溯到了乔姆斯基对语言学方法论的描述。在哲学史中,甚至在分析哲学的早期,这个术语都是找
> 102 不到的……这并不是说历史上的哲学家和 20 世纪初期的哲学家并没有类似的哲学的做法。他们确实有这样的做法,他们只是不用“直觉”一词来描述它们罢了。(第 2 页)

这里的看法是:哲学家**不用**直觉来说话这一事实并不意味着他们没有在**讨论**直觉,而且一旦我们弄清楚了这里所说的东西,那么很明显直觉**确实**在这些讨论中扮演着一个核心角色,即使这个名称没有出现。

当然,哲学直觉确实在当代哲学中扮演着一个重要角色这一事实并不意味着它们要么**应该**扮演这个角色,要么**需要**扮演这一角色。哲学家并不总想到他们自己那种方法论上的最大利益。蒂莫西·威廉姆森(2007 年)似乎采用了这一观点,他提出在方法论上哲学直觉是不重要的,同时,哲学事实上从自身与哲学直觉的疏离中得到了一些好处。威廉姆森相信哲学直觉在方

---

① 市川(即将出版)给出了相同的论点,提出我们很难得到那种哲学家在哲学讨论中明确诉诸我们的哲学直觉的例子,除非是那些哲学家(比如戈德曼,2007 年)在研究人类心理,即我们共同的或个人的哲学概念的例子中。

法论上是不重要的,因为他相信,至少在大部分例子中,我们有更好的证据。因此,比如,我们关于知识不仅仅是确证了的真信念的最好证据不是我们的如下哲学直觉,即"某人真正相信 p 这一点是被确证的,而不知道 p"是可能的,而是这是可能的这一事实。是这个直觉命题,而不是这个命题是直觉性的这一事实,被算作我们的最好的哲学证据。威廉姆森甚至走得更远,他不仅提出我们有着比哲学直觉更好的证据,还提出我们有充分的理由不把我们的哲学直觉视为证据,或者至少不把我们的哲学证据局限在哲学直觉之内,换言之,这样做①会招致哲学的怀疑主义。若将哲学证据局限在我们的哲学直觉之内,我们就承担上了这样一个责任,那就是要证明"一些命题对我们而言似乎是真的"为我们提供了"它是真的"的充分证据;威廉姆森认为我们无力承担这个责任。

简而言之,威廉姆森认为我们或许可以这样搞哲学,即不必诉诸哲学直觉来作为证据,而且实际上我们以这样的方式来搞 103
哲学会好很多,因为它避免了哲学怀疑主义。正如我在其他地方(亚历山大,2010 年)指明的,问题在于我们事实上不能搞那种我们想要搞的哲学,除非我们在某些情况下诉诸我们的哲学直觉。为了搞清楚原因,让我们再来考虑一下葛梯尔例子。根据威廉姆森的说法,我们关于知识不仅仅是确证了的真信念的最好证据是如下事实,即有可能某人有一个确证了的真信念 p 而不知道 p。因此,对于威廉姆森而言,如下这个论证既是有效的,又是正确的:

(1) 葛梯尔例子是可能的(也就是说,没有什么与这个

① 指的是把我们的哲学证据局限在哲学直觉之内。——译者

例子有关的东西是不一致的)。

(2) 如果一个葛梯尔例子出现,那么实验对象将会有一个确证了的真信念 p 而不知道 p。①

(3) 因此,有可能某人有一个确证了的真信念 p 而不知道 p。

(4) 因此,并不必然是只有当一个人有一个确证了的真信念 p 时,她才知道 p。

因为这个论证无论在什么地方都没有提及关于我们的心理事实,即我们的如下哲学直觉——有可能某人有一个确证了的真信念 p 而不知道 p,所以对威廉姆森来说似乎葛梯尔例子可以无须依赖我们的哲学直觉作为证据而成立。

但是,虽然这似乎是真的,即对(4)的论证并不需要一个前

---

① 实际上,威廉姆森是不是完全赞同这种讨论问题的方式,这一点也是不清楚的。威廉姆森首选的对葛梯尔论证的重构是这样的:

(A) 就像葛梯尔例子中的主角与相关的命题处在某种关系中那样,某人与某个命题 p 处在某种关系中,这一点也是可能的。

(B) 如果就像葛梯尔例子中的主角与相关的命题处在某种关系中那样,某人与某个命题 p 处在某种关系中,那么任何一个人,只要这个人像葛梯尔例子中的主角与相关的命题处在某种关系中那样与 p 处在某种关系中,这个人就有一个确证了的真信念 p,而 p 不是知识。

(C) 因此,某人有一个确证了的真信念,而这个信念不是知识,这一点是可能的。

尽管(A)和(C)与(1)和(3)是一样的,但是重要的是要知道(B)和(2)并不是一样的。实际上,威廉姆森否定了类似(2)的东西,其依据是存在一些补充葛梯尔思想实验的细节的方式,按照这种方式,这个小场景并不给出那种在其中确证了的真信念不是知识的例子。因为这一点是相当富有争议的,又因为没有什么东西能表明我们使用哪种表达式会带来什么结果,因此我们可以继续维持(1)到(4)。

提(这个前提报道了关于我们的心理事实,即我们的如下哲学直觉——有可能有人有一个确证了的真信念 p 而不知道 p),但是,这并不意味着我们的哲学直觉在哲学中没有扮演任何证据性的角色。哲学毕竟是一个**论证性**的实践,它并不仅仅涉及建立有效的论证,还要让人们(有时候是我们自身,有时候是其他人)信服那些论证是完好的(sound)。而且,不仅“为什么我们要接受(4)?”这个问题要出现——这个问题的答案是“因为(1)、(2)、(3)”。“为什么我们要接受(1)、(2)、(3)?”这个问题也要出现。这就是我们的哲学直觉似乎扮演着证据性角色的地方。

让我们看一下前提(2)。为什么我们要接受前提(2)呢? 104
葛梯尔依赖我们的“(2)为真”的哲学直觉作为证据。换言之,葛梯尔预计读者会接受(2),而且这种接受的基础也许就是对直觉的诉诸(读者的直觉,不仅是葛梯尔的直觉)。[1] 因此,直觉似乎**确实**在对(4)的论证中扮演着某种角色——尽管是以一种不直接的方式。它们并不为(4)的真本身提供证据。但是,它们确实为对(4)的论证所涉及的(至少一个)主要前提的真提供了证据。因此,虽然我们或许可以看到这个论证是有效的而无须诉诸我们的哲学直觉作为证据,那些直觉似乎确实扮演了一个角色,那就是向我们揭示这个论证是完好的。换言之,若不诉诸我们的哲学直觉,我们或许可以看到如果(1)、(2)、(3)是真的,那么(4)也是真的,而不能

---

① 重要的是要记住当哲学家诉诸直觉作为证据时,他们并不(至少不是典型地)只诉诸他们自己的直觉。哲学分析并不仅仅是思想自传,这一点是基于一些好理由的:只诉诸某人自己的直觉在论辩上没有什么效力。进一步讨论可见亚历山大和温伯格(2007 年)。

看到(4)是真的。①

在这一点上，似乎我们忽略了威廉姆森的“我们最好的哲学证据由关于世界的已知事实构成”这一看法。什么是我们接受(2)的基础呢？似乎是这样，根据威廉姆森的看法，我们接受(2)的基础就是如下**事实**，即如果一个葛梯尔例子出现了，那么这个对象将会有一个确证了的真信念 p 而不知道 p。② 但是，这会不会是正确的呢？从“如果一个葛梯尔例子出现了，那么这个对象就会有一个确证了的真信念 p 而不知道 p”这一事实到(2)的论证当然是有效的。毕竟，(2)的内容仅仅是“如果一个葛梯尔例子出现了，那么这个对象就会有一个确证了的真信念 p 而不知

---

① 市川(即将出版)介绍了一种针对不同种类的元哲学主张的有趣区分：

(1) 被直觉到的内容(通常)被认为是扶手椅哲学的重要证据/理由/材料/输入信号。

(2) 被直觉到的内容(通常)被认为是扶手椅哲学的重要证据/理由/材料/输入信号，因为它们是被直觉到的。

(3) 直觉状态，或关于直觉状态的事实(通常)被认为是扶手椅哲学的重要证据/理由/材料/输入信号。

在这些条目中，在这里被辩护的观点或者是(2)，或者是(2)和(3)的某种混合体。

② 在这里，我们假定证据的功能之一是为我们的信念提供一种认知上的基础。根据这种对证据的看法，如果威廉姆森是对的，即我们对于(2)的信念的证据由(2)这一事实构成，那么那个信念的认知基础仍然是(2)这一事实。尽管这并不是威廉姆森明确指派给证据的功能(可见威廉姆森，1997 年，2000 年)，但是他关于证据所说的话并没有排除它这方面的功能(可见威廉姆森，1997 年，第 728 页)。这样一来，就存在这样一些地方，在这里，对于我们的信念(2)，似乎威廉姆森有着另外的依据，即我们“运用认知概念的能力”(威廉姆森，2007 年，第 189 页)。然而，很难搞清楚这种看待事物的方式实际上是不是帮助了威廉姆森反对直觉的证据地位。毕竟，大多数参与了关于哲学直觉能否被视为证据的争论的哲学家都认为这种能力恰好就是产生认知**直觉**的能力。这将让(2)成为一个哲学直觉(至少在与所有关于直觉的证据地位的争论相关的意义上)，紧接着也就意味着直觉在对(4)的真的论证中扮演着一个直接的证据性角色。

道 p"这一主张。同时,任何命题都能够有效地推演出自身。问题在于那样一个论证将不会有任何的说服力。如果某人并不确信一个命题是真的,那么仅仅断言实际上它是这样的,那将会是于事无补的。因此,如果一个人并没有确信(2)是真的,那么断言实际上如果一个葛梯尔例子出现,那么这个对象就会有一个确证了的真信念 p 而不知道 p,这也将会是于事无补的。这样看来,如果一个葛梯尔例子出现,那么这个对象就会有一个确证了的真信念 p 而不知道 p 这一事实不能成为我们相信(2)的**基础**。

如果我们所说的是对的,那么似乎我们必须要么承认我们最好的哲学证据不仅由关于世界的已知事实构成,要么我们必须抛弃一种哲学的论辩观点,这种观点认为哲学是一种论证实践,其目的就是理性的说服,而且证据必须能够帮助达到这一目 105
的。面对着这两种选择,威廉姆森似乎愿意选择第二种,并抛弃了他称之为"证据的论辩标准"(dialectical standard of evidence)的东西。按照威廉姆森的说法,证据的论辩标准为怀疑主义留了门。[1] 如果采纳了证据必须能够说服那些未被说服的人这一看法,由于怀疑主义者是未被说服的,那么我们就采纳

---

① 除去这种关于证据的论辩标准的担忧外,威廉姆森还主张诉诸哲学的论辩本性并不会**支持**将哲学直觉视为证据的做法。根据威廉姆森的看法,存在着这样一些背景,在这里出现了(或者将会出现)关于什么才是真正的哲学直觉的争议,关于什么才被视为对哲学直觉的恰当表达的争议,以及关于为什么哲学直觉应该被视为证据的争议,还有是否存在像哲学直觉这样的东西的争议,等等。(比如,可见威廉姆森,2004 年,第 119—123 页;2007 年,第 235—237 页。)当然,即使威廉姆森是对的,可能存在一些我们无法在其中诉诸哲学直觉作为证据的背景这一事实并不意味着我们绝不能诉诸它们。为了支持这一更强的主张,威廉姆森将不得不证明这种问题在**所有**的背景中都存在。由于威廉姆森没有为这一更强的主张辩护,所以这似乎无法成为拒绝这样一幅哲学的论辩图画的理由,这幅图画视哲学直觉为证据。

了证据必须能够说服怀疑主义者这一看法。但是,因为怀疑主义者不仅仅是未被说服的,而且是不可能被说服的,所以这个观点只能通向失败——因为怀疑主义者不能被说服,证据的论辩标准设立了一个无法达到的标准。对威廉姆森来说,我们惟一的选择就是拒绝说服怀疑主义者的任务,并且一同拒绝证据的论辩标准。正如威廉姆森所言:

> 当人们有理由拒绝去玩怀疑主义者的论辩游戏时,证据的论辩标准就变得无足轻重了。在这种拒绝中,人们并没有抛弃他们对知识和理性的主张,因为恰当的证据标准是非论辩性的。就那个标准看来,怀疑主义者那种专横的质疑无法剥夺被质疑的事实作为证据的资格。忽略那些证据将会违背证据整体的要求。人们继续在证据的基础之上断定那种富有争议的命题,而不期望去寻找对这些命题的论证,这些论证仅仅使用那些被怀疑主义者所接受的前提和推理形式。由于从激进的怀疑主义困境中脱离出来是不可能的,因此首先人们必须高度谨慎,不要落入这种困境。(2004 年,第 123 页;2007 年,第 239 页)

然而,不清楚的是我们为什么不能接受这样一种证据理论,该理论要求证据能够说服那些未被说服的人,而不要求证据能够说服那些不能被说服的人。让我们在如下两个问题之间做出区分:“什么是你相信 p 为真的理由?”和“那是一个相信 p 为真的好(充足)理由吗?”怀疑主义者的目的是去质疑我们用肯定的命题来回答第二个问题的能力,不管我们对第一个问题的回答是什么。关于知觉的怀疑主义者很希望我们相信 p 的理由就是
106 我们的知觉 p,然而对有知觉 p 和 p 为真之间的关系提出怀疑。

关于记忆的怀疑主义者很希望我们相信 p 的理由就是我们的回忆 p,然而对回忆起 p 和 p 为真的关系提出怀疑。同时,关于判断的怀疑主义者很希望我们相信 p 的理由就是我们的直觉 p,然而对有直觉 p 和 p 为真的关系提出怀疑。没有被说服和不能被说服(即持怀疑主义的看法)之间的区别就在于一个人是不是愿意接受某种类型的理由作为相信 p 的足够好的理由。一个关于 p 的真仅仅未被说服的人很愿意承认存在一些足够好的理由来相信 p,并且期待听到这些理由;一个关于 p 的真不可能被说服的人不愿意承认这样的理由的存在。① 但是,由于没有被说服和不可能被说服之间的这一区别,我们似乎没有什么好的理由来认为为了让证据能够说服那些没有被说服的人,它必须也能够说服那些不可能被说服的人。这样一来,没有什么东西可以说服那些不可说服的人这一事实不能被用来反对这样一种证据理论,该理论需要证据能够说服那些未被说服的人。而且我们似乎没有理由不去接受一种不受怀疑主义的担忧所困扰的证据的论辩理论。

我们还有理由认为我们*应该*接受证据的论辩理论。考虑一下我们为什么重视证据。我们之所重视证据,是因为我们重视确证——换言之,相比那些没有被确证的信念,我们更加愿意拥有确证了的信念。为什么我们重视确证呢?一个自然而然的想法就是,我们之所以重视确证,是因为我们重视如下这一点:当我们的信念的真受到怀疑时,我们能够有能力为我们的信念辩

① 为了能够对付这样一些人——他们的怀疑主义态度是局部的而非整体的,他们仅仅否认某些种类的理由是相信 p 的足够好的理由,我们可能需要稍微调整一下这个修改了的论辩标准。然而重要之处在于我们能区分不被说服的和不能被(局部地或整体地)说服的人,这似乎是可能的,即去采用这样一种证据的论辩标准,这种标准不需要证据能够说服那些不能被说服的人。

护。我们经常参与那种需要我们为我们的信念辩护(不管是对我们自身还是对他人)的实践活动——换言之就是如同我们这样为信念提供理由,而这些理由能够说服那些未被说服的人——而且,我们认为能够成功地做到这一点也是重要的。因此,尽管一个人确证地相信 p 和能够确证她的信念 p 之间是有区别的,但是似乎我们之所以重视前者,至少部分地是因为我们也重视后者。现在,如果我们之所以重视证据是因为我们重视确证,而我们之所以重视确证,至少部分原因是我们重视我们确证我们的信念的能力,那么我们应该重视哪一些证据呢?看起
107 来,显而易见的答案是我们应该重视那些能够说服未被说服的人的证据——只有这种证据将会帮助我们确证我们的信念。这样一来,对为什么我们要重视证据的简单思考似乎揭示了为什么我们应该接受证据的论辩理论。

将这些观点放在一起,似乎我们有了接受证据的论辩理论的独立的理由,也就是说,证据的论辩理论解释了为什么我们首先要重视证据,而且我们可以接受证据的论辩理论同时避免招致激进怀疑主义的威胁。如果这是对的,那么鉴于我们要在承认哲学中的证据不仅仅由关于世界的已知事实构成和抛弃证据的论辩理论之间做出选择,似乎我们应该承认证据不仅仅由关于世界的已知事实构成。那么,结果是什么呢?我们先是担心在哲学上哲学直觉可能是不重要的。鉴于实验哲学和哲学直觉是重要的这一看法之间的紧密联系,对我们的哲学直觉的威胁也成为了对实验哲学的重要性的威胁。然而,最后我们发现,实际上我们不能去搞那种我们想要搞的哲学,除非我们在某些例子中诉诸我们的哲学直觉作为某些哲学主张为真的证据并作为相信它们为真的理由,同时诉诸这种心理证据不会招致哲学的怀疑主义。

5. 结 论

在结束对几个更为流行的为传统辩护的方式的考察后，我
们可以看到这些辩护有一个共同主题，那就是都认为实验哲学
家确实没有在研究那种正确的东西，这意味着实验哲学对我们
对传统哲学实践的理解没有带来什么贡献。但是，我们已经看
到，哲学直觉在做哲学家想做的至少一些事情时扮演着一个重
要角色，那种试图限定谁的直觉重要或哪一些直觉重要的尝试
带来一些悬而未决的经验问题，这些问题要求有更多的，而不是 108
更少的实验哲学。还有一点是值得注意的，因为每一种为传统
辩护的方法都涉及对一些思考哲学的相当普遍的方式的背离，
所以，我们发现自己处在一个略显矛盾的位置上，我们通过为这
些思考哲学的方式辩护来为实验哲学辩护，这样我们才能转回
去，重新质疑那些思考哲学的方式。

后　记

1. 在进行中……

我的目的是刻画出一场正在成长中的运动，关注的是实验 109
哲学和哲学直觉为我们提供重要哲学洞见这一信念之间的关
系——我们基于它们解释我们的哲学直觉的能力来提出哲学理
论，并且诉诸这些直觉作为那些理论为真的证据和相信这些理
论的理由。根据这种思考哲学的方式，对研究哲学认知感兴趣
是合情合理的，而且实验哲学是作为一种研究人们如何考虑哲
学问题的崭新方法而出现的，这种方法采用了社会科学和认知
科学的方法。它的结果是很有趣的，它揭示了一些关于我们心
灵的运作和我们思考哲学问题的方式的重要洞见。这些洞见又
反过来引发了有关哲学方法论的新问题，并且促使人们回过头
来思考我们在搞哲学时能够，并且应该采用哪些方法。当然，
要说的话还有很多。还有一些需要回答的问题，一些尚须获
得的细节，亟待发展的见解。这是对一场不断前进的运动的初 110
步研究，它已经改变了哲学的风貌，但是它的全景仍在展开
之中。

2. 关于实验哲学的方法的一个评注

实验哲学中的很大一部分工作涉及对调查方法的使用;人们被要求去阅读一些哲学小场景,并对其作出回应,这些回应被认为告诉了我们一些重要的东西,这些东西是与人们关于所说的东西的哲学直觉有关的。最近出现了关于调查的方法论的担忧。这里值得花一点时间来谈谈其中的两种担忧:一种与人们如何阅读哲学小场景有关;另一种则与调查回应和哲学直觉的关系有关。

可把握性和添加(Accessibility and Addition)

我们都了解哲学小场景中给出的信息并不是总能被读者所把握的,当读者在形成关于小场景中所说的东西的判断时,有时候他们会利用一些没有包含在小场景自身之中的信息。一些细节被忽略了,另外的一些则被添加了进来。根据恩斯特·索萨(2009年)的看法,这应该引起实验哲学家的关注。实验哲学的方法要求读者形成关于哲学小场景的判断。我们如何知道这些读者仅仅是在利用包含在他们被要求去考虑的那个小场景中的信息呢?索萨的担心是,如果我们无法确定读者是不是在对同样的信息作出回应,那么我们应该从对不一致的直觉的报道中得出什么结论这一点是不清楚的。也许直觉的不一致仅仅是表面的。如果不同的读者在利用不同的信息,那么那些似乎对相同的哲学小场景有着不同的直觉的读者也许只是对不同的哲学小场景有着不同的直觉(这是恰
111 当的)。索萨建议,除非我们可以排除这一点,否则,我们就

应该对实验哲学家在质疑传统哲学方法论时赋予直觉不一致的例子的重要性感到担忧。

然而,正如约书亚·亚历山大和乔纳森·温伯格(2007 年)所指出的,这一论证有两个问题。首先,这个论证依赖于一个悬而未决的经验假说,即不同的人对哲学小场景的解释有着**系统性的**不同,而且这一点在他们形成关于所说的东西的判断时扮演着一个重要的角色。然而,证明这种假说的惟一方法,就是进行对哲学认知的仔细研究,因此这个论证似乎要求有更多的,而不是更少的实验哲学。其次,这个论证对双方都起作用。无论这个论证为实验哲学带来什么问题,它也会为标准的哲学实践带来这些问题。这个论证揭示出,当人们在谈论具体的哲学小场景时,他们永远不能确定他们在谈论的是同一个东西。但是,如果这是真的,那么不仅我们从关于直觉不一致的报道中应该得出什么结论这一点是不清楚的,而且我们应该从关于直觉**一致性**的报道中得出什么结论这一点也是不清楚的。结果是某种直觉的怀疑主义,这种怀疑主义对于传统的哲学方法论和实验哲学来说都是成问题的。

调查回应和哲学直觉

实验哲学家要求人们去阅读一些哲学小场景并对其作出回应,但是人们对关于哲学小场景的问题的回答有时不仅反映了他们关于小场景本身的判断,还反映了这些回答敏感于会话语境(conversational context)和具体的会话规范(conversational norms)(史怀泽[Schwarz],1995 年,1996 年)。根据西蒙·库伦(Simon Cullen)的看法,这应该让实验哲学家感到担忧。实验哲学家让人们在一种奇怪的,有时是模棱两可的会话语境中阅读、思考那些哲学小场景。会话语境影响了人们解释这些小

场景和关于这些小场景他们被问到的问题的方式,会话的语境越独特,人们就越有可能依赖会话规范来指导他们的行为。库
112 伦相信,这让我们有了很好的理由来担心人们对那些与哲学小场景有关的问题的回应(库伦称之为**调查回应**)与他们的哲学直觉之间的关系。

库伦愿意承认实验哲学家已经揭示了调查回应展示了多样性和敏感性的一些有趣的模式。换言之,他愿意接受不同的人对有关具体哲学小场景的问题给出了不同的调查回应,而且那种有关具体哲学小场景的调查回应是随着比如会话语境这样的东西的变化而变化的。他想要提出的是,这并没有给我们任何理由来认为哲学直觉展示了多样性和敏感性的一些有趣模式。但是,重要的是要明白,这里有很多东西都取决于在我们看来哲学直觉应该是什么。如果我们认为哲学直觉不应该受到例如会话语境或会话规范的影响,那么担心调查回应并没有告诉我们一些与我们的哲学直觉有关的重要东西就是完全合理的了。问题在于库伦并没有给我们这样思考哲学直觉的理由,他只是采用了一个来自库皮尼(2007 年)的观点——根据后者的看法,哲学直觉并不敏感于实用的考虑——并且铺陈了这个观点给调查回应和哲学直觉的关系带来的影响。但是,还没有人给我们接受关于哲学直觉的这一具体看法的任何理由,也没有人给我们怀疑调查回应和哲学直觉之间的关系的任何理由。也许还会有担心调查回应和哲学直觉之间的关系的其他理由,但是,在缺乏对于哲学直觉并不敏感于实用的考虑这一点的论证的情况下,有关调查回应敏感于实用考虑的证据并不是它们没有揭示关于我们的哲学直觉的一些重要东西的证据。

最后还有一点值得在这里提出。尽管调查方法在实验哲学中仍然扮演着一个重要的角色,近些年来,那些实验哲学家用来

研究人们对哲学小场景作出回应的方式的统计方法和分析也已变得纯熟了起来(比如,对中介分析[mediation analyses]和结构方程模型[structural equation modeling]的运用)。实验哲学家还开始运用很多种来自社会科学和认知科学的方法:行为研 113
究、认知负荷的研究、眼球追踪研究、功能磁共振成像研究、反应时间研究,等等。对方法论的忧虑以后当然还会出现,正如它们在任何还在发展的经验性事业中都会出现。但是,我们应该如其所是地看待这些,换言之,这是改善我们在研究哲学认知时所使用的经验方法的机会。

3. 作为一场革命的实验哲学

实验哲学有时候被描述为一场革命运动,与更为传统的思考哲学的方式的一次重要的,也许甚至是剧烈的决裂。在某种意义上,这种刻画是正确的。一般说来,实验哲学所采用的方法与社会科学和认知科学联系更为紧密,实验哲学已经揭示了人们以一种很难预料的方式思考哲学问题,并且引发了有关我们中间的很多人搞哲学的方式的担忧。但是,重要的是要记住,在实验哲学家追问的那些问题之中没有什么革命性的东西。他们感兴趣的是一些关于世界和我们自身的相当传统的问题,关于我们如何思考哲学问题的问题,关于我们在试图回答这些问题时所采用的方法的问题,从西方哲学的最初源头苏格拉底的对话开始,这些问题就一直塑造着它。还有一点也很重要的,那就是要记住在哲学中有一个值得尊重的传统,这个传统用新的方法来质疑着关于人们如何思考哲学问题以及如何最佳地寻求这些问题的答案的传统预设。正如理查德・罗蒂(1967 年)在他为一本由早期哲学革命的一些最重要的著作所构成的合集的引

言的前几句话中写到的那样:“哲学的历史穿插着对先前哲学家的实践的反抗……”哲学史就是革命史,实验哲学只是它最近的一章。这一章仍在书写中。

参 考 文 献

Adams, F. and Steadman, A. 2004a: Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding? *Analysis* 64, 173–181.

Adams, F. and Steadman, A. 2004b: Intentional Action and Moral Considerations: Still Pragmatic. *Analysis* 64, 268–276.

Alexander, J. 2010: Is Experimental Philosophy Philosophically Significant? *Philosophical Psychology* 23, 377–389.

Alexander, J. and Weinberg, J. 2007: Analytic Epistemology and Experimental Philosophy. *Philosophy Compass* 2, 56–80.

Alexander, J., Mallon, R., and Weinberg, J. 2010a: Accentuate the Negative. *Review of Philosophy and Psychology* 1, 297–314.

Alexander, J., Mallon, R., and Weinberg, J. 2010b: Competence: What's In? What's Out? Who Knows? *Behavioral and Brain Sciences* 33, 329–330.

Alicke, M. 2008: Blaming Badly. *Journal of Cognition and Culture* 8, 179–186.

Annas, J. 2003: Virtue Ethics and Social Psychology. *A Priori* 2, 20–59.

Appiah, A. 2008: *Experiments in Ethics*. Cambridge, MA: Harvard University Press.

Arico, A. 2010: Folk Psychology, Consciousness, and Context Effects. *Review of Philosophy and Psychology* 1, 371–393.

Arico, A., Fiala, B., Goldberg, R., and Nichols, S. 2011: The Folk Psychology of Consciousness. *Mind & Language* 22, 21–35.

Baron, J. 2000: *Thinking and Deciding*, 3rd edn. New York: Cambridge University Press.

Bealer, G. 1992: The Incoherence of Empiricism. *Proceedings of the Aristotelian Society* 66, 99–138.

Bealer, G. 1996a: A Priori Knowledge and the Scope of Philosophy. *Philosophical Studies* 66, 121–142.

Bealer, G. 1996b: On the Possibility of Philosophical Knowledge. *Philosophical Perspectives* 10, 1–34.

Bealer, G. 1998: Intuition and the Autonomy of Philosophy. In M. DePaul and W. Ramsey (eds.) *Rethinking Intuition*. Lanham, MD: Rowman and Littlefield, 201–240.

Beebe, J. and Buckwalter, W. 2010: The epistemic side-effect effect. *Mind & Language* 25, 474–498.

Bengson, J. Forthcoming: Experimental Attacks on Intuitions and Answers. *Philosophy and Phenomenological Research.*

Berker, S. 2009: The Normative Insignificance of Neuroethics. *Philosophy & Public Affairs* 90, 188–209.

Boghossian, P. 2006: *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

BonJour, L. 1992: A Rationalist Manifesto, *Canadian Journal of Philosophy Supplementary Volume 18*, 53–88.

BonJour, L. 1998: *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.

Brown, J. Forthcoming: Experimental Philosophy, Contextualism and SSI. *Philosophy and Phenomenological Research*, doi: 10.1111/j.1933-1592.2010.00461.x

Buckwalter, W. 2010: Knowledge Isn't Closed on Saturdays. *Review of Philosophy and Psychology* 1, 395–406.

Buckwalter, W. Forthcoming: The Mystery of Stakes and Error in Ascriber Intuitions. Manuscript under review.

Buckwalter, W. and Stich, S. Forthcoming: Gender and Philosophical Intuitions. Manuscript under review.

Cacioppo, J.T. and Petty, R.E. 1982: The Need for Cognition. *Journal of Personality and Social Psychology* 42, 116–131.

Camp, E. 2009: Two Varieties of Literary Imagination: Metaphor, Fiction, and Thought Experiments. *Midwest Studies in Philosophy* 33, 107–130.

Chalmers, D. 1996: *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chomsky, N. 1965: *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.

Christensen, D. 2007: Epistemology and Disagreement: The Good News. *The Philosophical Review* 116, 187–217.

Cohen, S. 1998: Contextualist Solutions to Epistemological Problems:

Skepticism, Gettier, and the Lottery. *Australasian Journal of Philosophy* 76, 289–306.

Cokely, E. and Feltz, A. 2009: Individual Differences, Judgment Biases, and Theory-of-Mind: Deconstructing the Intentional Action Side Effect Asymmetry. *Journal of Research in Personality* 43, 18–24.

Cullen, S. 2010: Survey-Driven Romanticism. *Review of Philosophy and Psychology* 1, 275–296.

Cummins, R. 1998: Reflection on Reflective Equilibrium. In M. DePaul and W. Ramsey (eds.) *Rethinking Intuition*. Lanham, MD: Rowman and Littlefield, 113–128.

Cushman, F. and Mele, A. 2008: Intentional Action: Two-and-a-Half Folk Concepts? In J. Knobe and S. Nichols (eds.) *Experimental Philosophy*. Oxford: Oxford University Press, 171–188.

Darley, J. and Batson, C. 1973: From Jerusalem to Jericho: A Study of Situational and Dispositional Variables in Helping Behavior. *Journal of Personality and Social Psychology* 27, 100–108.

DeRose, K. 1992: Contextualism and Knowledge Attributions. *Philosophy and Phenomenological Research* 52, 172–198.

DeRose, K. 1995: Solving the Skeptical Problem. *Philosophical Review* 104, 1–52.

DeRose, K. 2011: Contextualism, Contrastivism, and X-Phi Surveys. *Philosophical Studies* 156, 81–110.

Deutsch, M. 2009: Experimental Philosophy and the Theory of Reference. *Mind & Language* 24, 445–466.

Deutsch, M. 2010: Intuitions, Counter-Examples, and Experimental Philosophy. *Review of Philosophy and Psychology* 1, 447–460.

Devitt, M. 1994: The Methodology of Naturalistic Semantics. *The Journal of Philosophy* 91, 545–572.

Devitt, M. 2006: Intuitions in Linguistics. *British Journal for the Philosophy of Science* 57, 481–513.

Devitt, M. 2010: Experimental Semantics. *Philosophy and Phenomenological Research* 82, 418–435.

Doris, J. 1998: Persons, Situations, and Virtue Ethics. *Nous* 32, 504–530.

Doris, J. 2002: *Lack of Character*. Cambridge: Cambridge University Press.

Earlenbaugh, J. and Molyneux, B. 2009a: If Intuitions Must Be Evidential Then Philosophy Is in Big Trouble. *Studia Philosophica Estonica* 2, 35–53.

Earlenbaugh, J. and Molyneux, B. 2009b: Intuitions Are Inclinations to Believe. *Philosophical Studies* 145, 89–109.

Elga, A. 2006: Reflection and Disagreement. *Nous* 41, 478–502.

Elgin, C. 1996: *Considered Judgment*. Princeton, NJ: Princeton University Press.

Ericsson, K., Charness, N., Feltovich, P., and Hoffman, R. (eds.) 2006:

The Cambridge Handbook of Expertise and Expert Performance. Cambridge: Cambridge University Press.

Fantl, J. and McGrath, M. 2002: Evidence, Pragmatics, and Justification. *The Philosophical Review* 111, 67–94.

Feldman, R. 2006: Epistemological Puzzles about Disagreement. In S. Heatherington (ed.) *Epistemology Futures*. Oxford: Oxford University Press, 216–235.

Feldman, R. and Warfield, F. 2007: *Disagreement*. Oxford: Oxford University Press.

Feltz, A. and Cokely, E. 2007: An Anomaly in Intentional Action Ascription: More Evidence of Folk Diversity. In D. McNamara and J. Trafton (eds.) *Proceedings of the 29th Annual Cognitive Science Society*. Austin, TX: Cognitive Science Society, 1748.

Feltz, A. and Cokely, E. 2009: Do Judgments About Freedom and Responsibility Depend On Who You Are? Personality Differences in Intuitions about Compatibilism and Incompatibilism. *Consciousness and Cognition* 18, 342–350.

Feltz, A. and Zarpentine, C. 2010: Do You Know More When It Matters Less? *Philosophical Psychology* 23, 683–706.

Feltz, A., Cokely, E., and Nadelhoffer, T. 2009: Natural Compatibilism versus Natural Incompatibilism: Back to the Drawing Board. *Mind & Language* 24, 1–23.

Fields, J. and Schuman, H. 1976: Public Beliefs about the Beliefs of the Public. *Public Opinion Quarterly* 40, 427–488.

Foot, P. 1967: The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review* 5, 5–15.

Frankfurt, H. 1969: Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy* 66, 829–839.

Gendler, T. Szabó, 2007: Philosophical Thought Experiments, Intuitions, and Cognitive Equilibrium. *Midwest Studies in Philosophy* 31, 68–89.

Gettier, E. 1963: Is Justified True Belief Knowledge? *Analysis* 23, 121–123.

Glass, A., Holyoak, K., and Santa, J. 1979: *Cognition*. Reading, MA: Addison-Wesley.

Goldman, A. 2007: Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status. *Grazer Philosophische Studien* 74, 1–26.

Goldman, A. and Pust, J. 1998: Philosophical Theory and Intuitional Evidence. In M. DePaul and W. Ramsey (eds.) *Rethinking Intuition*. Lanham, MD: Rowman and Littlefield, 179–200.

Gray, H., Gray, K., and Wegner, D. 2007: Dimensions of Mind Perception. *Science* 315, 619.

Gray, K. and Wegner, D. 2009: Moral Typecasting. *Journal of Personality and Social Psychology* 96, 505–520.

Greene, J. 2003: From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology? *Nature Reviews Neuroscience* 4, 846–850.

Greene, J. 2008: The Secret Joke of Kant's Soul. In W. Sinnott-Armstrong (ed.) *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, MA: MIT Press, 35–80.

Greene, J., Sommerville, R., Nystrom, L., Darley, J., and Cohen, J. 2001: An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science* 293, 2105–2108.

Grice, P. 1989: *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Griffiths, P. and Stotz, K. 2006: Genes in the Postgenomic Era. *Theoretical Medicine and Bioethics* 27, 499–521.

Griffiths, P. and Stotz, K. 2007: Gene. In D. Hull and M. Ruse (eds.) *Cambridge Companion to the Philosophy of Biology*. Cambridge: Cambridge University Press, 85–102.

Grundmann, T. 2010: Some Hope for Intuitions: A Reply to Weinberg. *Philosophical Psychology* 23, 481–509.

Guglielmo, S. and Malle, B. 2009: The Timing of Blame and Intentionality: Testing the Moral Bias Hypothesis. Poster presented at the Annual Meeting of the Society for Philosophy and Psychology, Bloomington, IN.

Guglielmo, S. and Malle, B. 2010: Can Unintended Side-Effects Be Intentional? Solving a Puzzle in People's Judgments of Intentionality and Morality. *Personality and Social Psychology Bulletin* 36, 1635–1647.

Haidt, J. 2001: The Emotional Dog and Its Rationalist Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review* 108, 814–834.

Haidt, J., Koller, S., and Dias, M. 1993: Affect, Culture and Morality. *Journal of Personality and Social Psychology* 65, 613–628.

Hales, S. 2000: The Problem of Intuition. *American Philosophical Quarterly* 37, 135–147.

Hales, S. 2006: *Relativism and the Foundations of Philosophy*. Cambridge, MA: MIT Press.

Harman, G. 1999: Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society* 99, 315–331.

Hawthorne, J. 2004: *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hawthorne, J. and Stanley, J. 2008: Knowledge and Action. *The Journal of Philosophy* 105, 571–590.

Henderson, D. and Horgan, T. 2000: What Is the A Priori and What Is It Good For? *Southern Journal of Philosophy* 38, 51–86.

Henderson, D. and Horgan, T. 2001: The A Priori Isn't All It Is Cracked Up To Be, But It Is Something. *Philosophical Topics* 29, 219–250.

Hintikka, J. 1999: The Emperor's New Intuitions. *The Journal of Philosophy* 96, 127–147.

Hitchcock, C. and Knobe, J. 2009: Cause and Norm. *The Journal of Philosophy* 106, 587–612.

Horvath, J. 2010: How (Not) to React to Experimental Philosophy. *Philosophical Psychology* 23, 447–480.

Huebner, B., Bruno, M., and Sarkissian, H. 2010: What Does the Nation of China Think about Phenomenal Consciousness? *Review of Philosophy and Psychology* 1, 225–243.

Ichikawa, J. Forthcoming: Who Needs Intuitions? Two Experimentalist Critiques. In A. Booth and D. Rowbottom (eds.) *Intuitions*. Oxford: Oxford University Press.

Ichikawa, J. and Jarvis, B. 2009: Thought-Experiment Intuitions and Truth in Fiction. *Philosophical Studies* 142: 221–246.

Ichikawa, J., Maitra, I., and Weatherson, B. Forthcoming: In Defense of a Kripkean Dogma. *Philosophy and Phenomenological Research*, doi: 10.1111/j.1933-1592.2010.00478.x.

Isen, A. and Levin, P. 1972: Effects of Feeling Good on Helping: Cookies and Kindness. *Journal of Personality and Social Psychology* 21, 384–388.

Jackman, H. 2001: Ordinary Language, Conventionalism and *a priori* Knowledge. *Dialectica* 55, 315–325.

Jackson, F. 1998a: *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.

Jackson, F. 1998b: Reference and Description Revisited. *Philosophical Perspectives* 12, 201–218.

Kamm, F. 1989: Harming Some to Save Others. *Philosophical Studies* 57, 227–60.

Kamm, F. 2007: *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford: Oxford University Press.

Kamtekar, R. 2004: Situationism and Virtue Ethics on the Content of Our Character. *Ethics* 114, 458–491.

Kane, R. 1999: Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism. *The Journal of Philosophy* 96, 217–240.

Kauppinen, A. 2007: The Rise and Fall of Experimental Philosophy. *Philosophical Explorations* 10, 95–118.

Kelly, T. 2005: The Epistemic Significance of Disagreement. In J. Hawthorne and T. Gendler Szabó (eds.) *Oxford Studies in Epistemology*, Vol. 1. Oxford: Oxford University Press, 167–196.

Kelly, T. 2007: Peer Disagreements and Higher Order Evidence. In R. Feldman and T. Warfield (eds.) *Disagreement*. Oxford: Oxford University Press, 111–174.

Kelly, T. 2008: Disagreement, Dogmatism, and Belief Polarization. *The Journal of Philosophy* 105: 611–633.

Knobe, J. 2003a: Intentional Action and Side-effects in Ordinary Language. *Analysis* 63, 190–194.

Knobe, J. 2003b: Intentional Action in Folk Psychology: An Experimental Investigation. *Philosophical Psychology* 16, 309–323.

Knobe, J. 2004a: Folk Psychology and Folk Morality: Response to Critics. *Journal of Theoretical and Philosophical Psychology* 24, 270–279.

Knobe, J. 2004b: Intention, Intentional Action, and Moral Considerations. *Analysis* 64, 181–187.

Knobe, J. 2006: The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology. *Philosophical Studies* 130, 203–231.

Knobe, J. 2007a: Experimental Philosophy. *Philosophy Compass* 2, 81–92.

Knobe, J. 2007b: Experimental Philosophy and Philosophical Significance. *Philosophical Explorations* 10, 119–122.

Knobe, J. and Burra, A. 2006: Intention and Intentional Action: A Cross-Cultural Study. *Journal of Culture and Cognition* 6, 113–132.

Knobe, J. and Kelly, S. 2009: Can One Act for a Reason Without Acting Intentionally? In C. Sandis (ed.) *New Essays on the Explanation of Action*. Basingstoke: Palgrave Macmillan, 169–183.

Knobe, J. and Mendlow, G. 2004: The Good, the Bad and the Blameworthy: Understanding the Role of Evaluative Reasoning in Folk Psychology. *Journal of Theoretical and Philosophical Psychology* 24, 252–258.

Knobe, J. and Nichols, S. 2008. An Experimental Philosophy Manifesto. In J. Knobe and S. Nichols (eds.) *Experimental Philosophy*. Oxford: Oxford University Press, 3–16.

Knobe, J. and Prinz, J. 2008: Intuitions about Consciousness. *Phenomenology and the Cognitive Sciences* 1, 67–83.

Kornblith, H. 1998: The Role of Intuitions in Philosophical Inquiry: An Account with No Unnatural Ingredients. In M. DePaul and W. Ramsey (eds.) *Rethinking Intuition*. Lanham, MD: Rowman and Littlefield, 129–142.

Kornblith, H. 2002: *Knowledge and Its Place in Nature*. Oxford: Oxford University Press.

Kornblith, H. 2007: Naturalism and Intuitions. *Grazer Philosophische Studien* 72, 27–49.

Kornblith, H. 2010: What Reflective Endorsement Cannot Do. *Philosophy and Phenomenological Research* 80, 1–19.

Kripke, S. 1980: *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Kupperman, J. 2001: The Indispensability of Character. *Philosophy* 76, 239–250.

Lehrer, K. 1990: *Theory of Knowledge*. Oxford: Routledge.

Leslie, A., Knobe, J., and Cohen, A. 2006: Acting Intentionally and the Side-effect Effect: Theory of Mind and Moral Judgment. *Psychological Science* 17, 421–427.

Levin, J. 2005: The Evidential Status of Philosophical Intuition. *Philosophical Studies* 121, 193–224.

Lewis, D. 1978: Truth in Fiction. *American Philosophical Quarterly* 15, 37–46. Reprinted in D. Lewis, *Philosophical Papers: Volume I*. New York: Oxford University Press.

Lewis, D. 1983: *Philosophical Papers: Volume I*. New York: Oxford University Press.

Lewis, D. 1996: Elusive Knowledge. *Australasian Journal of Philosophy* 74, 549–567.

Liao, S.M. 2008: A Defense of Intuitions. *Philosophical Studies* 140, 247–262.

Liao, S.M., Wiegmann, A., Alexander, J., and Vong, G. Forthcoming: Putting the Trolley in Order: Experimental Philosophy and the Loop Case. *Philosophical Psychology*.

Livengood, J., Sytsma, J., Feltz, A., Scheines, R., and Machery, E. 2010: Philosophical Temperament. *Philosophical Psychology* 23, 313–330.

Lombrozo, T. 2006: The Structure and Function of Explanations. *Trends in Cognitive Sciences* 10, 464–470.

Lombrozo, T. 2011: The Instrumental Value of Explanations. *Philosophy Compass* 6, 539–551.

Lombrozo, T. and Carey, S. 2006: Functional Explanation and the Function of Explanation. *Cognition* 99, 167–204.

Ludwig, K. 2007: The Epistemology of Thought Experiments: First Person versus Third Person Approaches. *Midwest Studies in Philosophy* 31, 128–159.

Ludwig, K. 2010: Intuitions and Relativity. *Philosophical Psychology*, 23, 427–445.

Lynch, M. 2006: Trusting Intuitions. In P. Greenough and M. Lynch (eds.) *Truth and Realism*. Oxford: Oxford University Press, 227–238.

Machery, E. 2008: Understanding the Folk Concept of Intentional Action: Philosophical and Experimental Issues. *Mind & Language* 2, 165–189.

Machery, E. 2010: The Bleak Implications of Moral Psychology. *Neuroethics* 3, 223–231.

Machery, E., Mallon, R., Nichols, S., and Stich, S. 2004: Semantics, Cross-Cultural Style. *Cognition* 92, B1–B12.

Malle, B. 2006: Intentionality, Morality, and their Relationship in Human Judgment. *Journal of Cognition and Culture* 6, 87–112.

Malle, B. and Nelson, S. 2003: Judging *mens rea:* The Tension between Folk Concepts and Legal Concepts of Intentionality. *Behavioral Sciences and the Law* 21, 563–580.

Mallon, R. 2009: Knobe vs. Machery: Testing the Trade-Off Hypothesis. *Mind & Language* 23, 246–255.

Mallon, R., Machery, E., Nichols, S., and Stich, S. 2009: Against Arguments from Reference. *Philosophy and Phenomenological Research* 79, 332–356.

Marr, D. 1982: *Vision: A Computational Approach*. San Francisco, CA: Freeman & Co.

Marti, G. 2009: Against Semantic Multiculturalism. *Analysis* 69, 42–48.

May, J., Sinnott-Armstrong, W., Hull, J.G., and Zimmerman, A. 2010: Practical Interests, Relevant Alternatives, and Knowledge Attributions: An Empirical Study. *Review of Philosophy and Psychology* 1, 265–273.

McCann, H. 2005: Intentional Action and Intending: Recent Empirical Studies. *Philosophical Psychology* 18: 737–748.

Mele, A. 2001: Acting Intentionally: Probing Folk Notions. In B. Malle, L. Moses, and D. Baldwin (eds.) *Intentions and Intentionality: Foundations of Social Cognition*. Cambridge, MA: MIT Press, 27–44.

Mele, A. and Cushman, F. 2007: Intentional Action, Folk Judgments, and Stories: Sorting Things Out. *Midwest Studies in Philosophy* 31, 184–201.

Milgram, S. 1974: *Obedience to Authority: An Experimental View*. New York: Harper & Row.

Monin, B. and Miller, D. 2001: Moral Credentials and the Expression of Prejudice. *Journal of Personality and Social Psychology* 81, 33–43.

Nadelhoffer, T. 2004a: On Praise, Side Effects, and Folk Ascriptions of Intentionality. *Journal of Theoretical and Philosophical Psychology* 24, 196–213.

Nadelhoffer, T. 2004b: Blame, Badness, and Intentional Action: A Reply to Knobe and Mendlow. *Journal of Theoretical and Philosophical Psychology* 24, 259–269.

Nadelhoffer, T. 2005: Skill, Luck, Control, and Intentional Action. *Philosophical Psychology* 18, 341–352.

Nadelhoffer, T. 2006a: Desire, Foresight, Intentions, and Intentional Action: Probing Folk Intuitions. *Journal of Cognition and Culture* 6, 133–157.

Nadelhoffer, T. 2006b: Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Juror Impartiality. *Philosophical Explorations* 9, 203–220.

Nadelhoffer, T. and Nahmias, E. 2007: The Past and Future of Experimental Philosophy. *Philosophical Explorations* 10, 123–149.

Nado, J. 2008: Effects of Moral Cognition on Judgments of Intentionality. *British Journal for the Philosophy of Science* 59, 709–731.

Nagel, J. 2007: Epistemic Intuitions. *Philosophy Compass* 2, 792–819.

Nagel, J. 2010: Knowledge Ascriptions and the Psychological Consequences of Thinking about Error. *The Philosophical Quarterly* 60, 286–306.

Nagel, J. Forthcoming: Intuitions and Experiments: A Defense of the Case Method in Epistemology. *Philosophy and Phenomenological Research*.

Nahmias, E. and Murray, D. 2011: Experimental Philosophy on Free Will: An Error Theory for Incompatibilist Intuitions. In J. Aguilar, A. Buckareff, and K. Frankish (eds.) *New Waves in Philosophy of Action*. New York: Palgrave Macmillan, 189–216.

Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J. 2004: The Phenomenology of Free Will. *Journal of Consciousness Studies* 11, 162–179.

Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J. 2005: Surveying Freedom: Folk Intuitions About Free Will and Moral Responsibility. *Philosophical Psychology* 18, 561–584.

Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J. 2006: Is Incompatibilism Intuitive? *Philosophy and Phenomenological Research* 73, 28–53.

Nahmias, E., Coates, D., and Kvaran, T. 2007: Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions. *Midwest Studies in Philosophy* 31, 214–242.

Nichols, S. 2004a: Folk Concepts and Intuitions: From Philosophy to Cognitive Science. *Trends in Cognitive Sciences* 8, 514–518.

Nichols, S. 2004b: *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford: Oxford University Press.

Nichols, S. and Knobe, J. 2007: Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions. *Nous*, 41, 663–685.

Nichols, S. and Ulatowski, J. 2007: Intuitions and Individual Differences: The Knobe Effect Revisited. *Mind & Language* 22, 346–365.

Nichols, S., Stich, S., and Weinberg, J. 2003: Metaskepticism: Meditations in Ethno-Epistemology. In S. Luper (ed.) *The Skeptics: Contemporary Debates*. Burlington, VT: Ashgate Press, 227–248.

Nickerson, R. 1998: Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology* 2, 175–220.

Nisbett, R. 2003: *The Geography of Thought*. New York: Free Press.

Nisbett, R. and Norenzayan, A. 2002: Culture and Cognition. In H. Plasher, S. Yantis, D. Medin, R. Gallistel, and J. Waxted (eds.) *Steven's Handbook of Experimental Psychology*, 3rd edn. Hoboken, NJ: Wiley, 561–597.

Nisbett, R. and Wilson, T. 1977: Telling More than We Can Know: Verbal Reports on Mental Processes. *Psychological Review* 84, 231–259.

Nisbett, R., Peng, K., Choi, I., and Norenzayan, A. 2001: Culture and Systems of Thought: Holistic vs. Analytic Cognition. *Psychological Review* 108, 291–310.
Norton, M., Vandello, J., and Darley, J. 2004: Casuistry and Social Category Bias. *Journal of Personality and Social Psychology* 87, 817–831.
O'Connor, T. 2000: *Persons and Causes: The Metaphysics of Free Will.* New York: Oxford University Press.
Pereboom, D. 2001: *Living Without Free Will.* Cambridge: Cambridge University Press.
Petrinovich, L. and O'Neill, P. 1996: Influence of Wording and Framing Efects on Moral Intuitions. *Ethology and Sociobiology* 17, 145–171.
Pettit, D. and Knobe, J. 2009: The Pervasive Impact of Moral Judgment. *Mind & Language* 24: 586–604.
Phelan, M. Forthcoming: Evidence That Stakes Don't Matter For Evidence. Manuscript under review.
Phelan, M. and Sarkissian, H. 2009: Is the Trade-off Hypothesis Worth Trading For? *Mind & Language* 24, 164–180.
Pinillos, N.A. Forthcoming: Knowledge, Experiments, and Practical Interests. In J. Brown and M. Gerken (eds.) *New Essays On Knowledge Ascriptions.* Oxford: Oxford University Press.
Pinillos, N.A. and Simpson, S. Forthcoming: Experimental Evidence in Support of Anti-Intellectualism about Knowledge. Manuscript under review.
Pink, T. 2004: *Free Will: A Very Short Introduction.* New York: Oxford University Press.
Plant, E. and Devine, P. 1998: Internal and External Sources of Motivation to Respond without Prejudice. *Journal of Personality and Social Psychology* 75, 811–832.
Plantinga, A. 1993: *Warrant and Proper Function.* Oxford: Oxford University Press.
Prinz, J. 2007: *The Emotional Construction of Morals.* Oxford: Oxford University Press.
Pronin, E., Lin, D., and Ross, L. 2002: The Bias Blindspot: Perceptions of Bias in Self Versus Others. *Personality and Social Psychology Bulletin* 28, 369–381.
Pust, J. 2000: *Intuitions as Evidence.* New York: Garland Publishing.
Pylyshyn, Z. 1984: *Computation and Cognition.* Cambridge, MA: MIT Press.
Robbins, P. and Jack, A. 2006: The Phenomenal Stance. *Philosophical Studies* 127, 59–85.
Rorty, R. 1967: *The Linguistic Turn.* Chicago, IL: University of Chicago Press.

Roskies, A. and Nichols, S. 2008: Bringing Moral Responsibility Down To Earth. *The Journal of Philosophy* 105, 371–388.

Ross, L. 1977: The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes. *Journal of Experimental Social Psychology* 13, 279–301.

Roxborough, C. and Cumby, J. 2009: Folk Psychological Concepts: Causation. *Philosophical Psychology* 22, 205–213.

Schaffer, J. 2005: Contrastive Knowledge. *Oxford Studies in Epistemology* 1, 235–271.

Schaffer, J. and Knobe, J. 2011: Contrastive Knowledge Surveyed. *Nous*, doi: 10.1111/j.1468-0068.2010.00795.x.

Schultz, E., Cokely, E., and Feltz, A. Forthcoming: Persistent Bias in Expert Judgments about Free Will and Moral Responsibility: A Test of the Expertise Defense. *Consciousness and Cognition*.

Schwarz, N. 1995: What Respondents Learn from Questionnaires: The Survey Interview and the Logic of Conversation. *International Statistical Review/Revue Internationale de Statistique* 63, 153–168.

Schwarz, N. 1996: *Cognition and Communication: Judgment Biases, Research Methods, and the Logic of Conversation*. Mahwah, NJ: Erlbaum.

Schwitzgebel, E. 2009: Do Ethicists Steal More Books? *Philosophical Psychology* 22, 711–725.

Schwitzgebel, E. and Cushman, F. 2012: Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind & Language*, forthcoming.

Schwitzgebel, E. and Rust, J. 2009: The Moral Behavior of Ethicists: Peer Opinion. *Mind* 118, 1043–1059.

Schwitzgebel, E. and Rust, J. 2010: Do Ethicists and Political Philosophers Vote More Often Than Other Professors? *Review of Philosophy and Psychology* 1, 189–199.

Schwitzgebel, E. and Rust, J. Forthcoming: The Self-Reported Behavior of Ethics Professors. Manuscript under review.

Schwitzgebel, E., Rust, J., Huang, L., Moore, A., and Coates, J. Forthcoming: Ethicists' Courtesy at Philosophy Conferences. *Philosophical Psychology*.

Shanteau, J. 1992: Competence in Experts: The Role of Task Characteristics. *Organizational Behavior and Human Decision Processes* 53, 252–266.

Singer, P. 1972: Moral Experts. *Analysis* 32, 115–117.

Singer, P. 1982: How Do We Decide? *The Hastings Center Report* 12, 9–11.

Singer, P. 2005: Ethics and Intuitions. *Journal of Ethics* 9, 331–352.

Sinnott-Armstrong, W. 2008: Framing Moral Intuitions. In W. Sinnott-Armstrong (ed.) *Moral Psychology, Volume 2: The Cognitive Science of Morality*. Cambridge, MA: MIT Press, 47–76.

Sommers, T. 2010: Experimental Philosophy and Free Will. *Philosophy Compass* 5, 199–212.
Sommers, T. 2011: In Memoriam: The X-Phi Debate. *The Philosophers' Magazine* 52, 89–93.
Sosa, E. 1991: *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
Sosa, E. 1996: Rational Intuition: Bealer on Its Nature and Epistemic Status. *Philosophical Studies* 81, 151–162.
Sosa, E. 1998: Minimal Intuition. In M. DePaul and W. Ramsey (eds.) *Rethinking Intuition*. Lanham, MD: Rowman and Littlefield, 257–270.
Sosa, S. 2007a: Experimental Philosophy and Philosophical Intuition. *Philosophical Studies* 132, 99–107.
Sosa, E. 2007b: Intuitions: Their Nature and Epistemic Efficacy. *Grazer Philosophische Studien* 74, 51–67.
Sosa, E. 2009: A Defense of the Use of Intuitions in Philosophy. In M. Bishop and D. Murphy (eds.) *Stich and His Critics*. Malden, MA: Wiley-Blackwell, 101–112.
Sreenivasan, G. 2002: Errors about Errors: Virtue Theory and Trait Attribution. *Mind* 111, 47–68.
Stalnaker, R. 1999: *Context and Content*, Oxford: Oxford University Press.
Stanley, J. 2005: *Knowledge and Practical Interests*. Oxford: Oxford University Press.
Starmans, C. and Friedman, O. Forthcoming: A Sex Difference in Adults' Attributions of Knowledge. Manuscript under review.
Stich, S. 1988: Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity. *Synthese* 74, 391–413.
Stich, S. 1990: *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA: MIT Press.
Stich, S. and Buckwalter, W. 2011: Gender and the Philosophy Club. *The Philosopher's Magazine* 52, 60–65.
Stich, S. and Nisbett, R. 1980: Justification and the Psychology of Human Reasoning. *Philosophy of Science* 47, 188–202.
Strawson, G. 1986: *Freedom and Belief*. New York: Oxford University Press.
Swain, S., Alexander, J., and Weinberg, J. 2008: The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp. *Philosophy and Phenomenological Research* 76, 138–155.
Sytsma, J. and Livengood, J. 2011: A New Perspective Concerning Experiments on Semantic Intuitions. *Australasian Journal of Philosophy* 89, 315–332.
Sytsma, J. and Machery, E. 2010: Two Conceptions of Subjective Experience. *Philosophical Studies* 151, 299–327.

Talbot, B. 2009: Psychology and the Use of Intuitions in Philosophy. *Studia Philosophica Estonia* 2, 157–176.

Thomson, J. 1985: The Trolley Problem. *The Yale Law Journal* 94, 1395–1415.

Tversky, A. and Kahneman, D. 1974: Judgment under Uncertainty: Heuristics and Biases. *Science* 185, 1124–1131. Reprinted in D. Kahneman, P. Slovic, and A. Tversiky (eds.) *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Uhlmann, E., Pizarro, D., Tannenbaum, D., and Ditto, P. 2009: The Motivated Use of Moral Principles. *Judgment and Decision Making* 4, 476–491.

van Inwagen, P. 1997: Materialism and the Psychological-Continuity Account of Personal Identity. *Nous* 31, 305–319.

Wason, P. 1960: On the Failure to Eliminate Hypotheses in a Conceptual Task. *The Quarterly Journal of Experimental Psychology* 12, 129–140.

Weatherson, B. 2003: What Good Are Counterexamples? *Philosophical Studies* 115, 1–31.

Weatherson, B. Forthcoming: Defending Interest-Relative Invariantism. Manuscript under review.

Weinberg, J. 2007: How To Challenge Intuitions Empirically Without Risking Skepticism. *Midwest Studies in Philosophy* 31, 318–343.

Weinberg, J. and Alexander, J. Forthcoming: The Challenge of Sticking with Intuitions through Thick and Thin. In A. Booth and D. Rowbottom (eds.) *Intuitions*. Oxford: Oxford University Press.

Weinberg, J., Nichols, S., and Stich, S. 2001: Normativity and Epistemic Intuitions. *Philosophical Topics* 29, 429–460.

Weinberg, J., Gonnerman, C., Buckner, C., and Alexander, J. 2010: Are Philosophers Expert Intuiters? *Philosophical Psychology* 23, 331–355.

Weinberg, J., Alexander, J., Gonnerman, C., and Reuter, S. Forthcoming: Restrictionism & Reflection: Challenge Deflected, or Simply Redirected? *The Monist*.

White, R. 2005. Epistemic Permissiveness. *Philosophical Perspectives* 19, 445–459.

Williamson, T. 1997: Knowledge as Evidence. *Mind* 106, 717–741.

Williamson, T. 2000: *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, T. 2004: Philosophical 'Intuitions' and Scepticism about Judgment. *Dialectica* 58, 109–155.

Williamson, T. 2005: Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking. *Proceedings of the Aristotelian Society* 105, 1–23.

Williamson, T. 2007: *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.

Williamson, T. 2011: Philosophical Expertise and the Burden of Proof. *Metaphilosophy* 42, 215–229.
Wilson, T. 2002: *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.
Wright, J. 2010: On Intuitional Stability: The Clear, the Strong, and the Paradigmatic. *Cognition* 115, 419–503.
Young, L., Cushman, F., Adolphs, R., Tranel, D., and Hauser, M. 2006: Does Emotion Mediate the Effect of an Action's Moral Status on Its Intentional Status? Neuropsychological Evidence. *Journal of Cognition and Culture* 6, 291–304.
Zamzow, J. and Nichols, S. 2009: Variations in Ethical Intuitions. *Philosophical Issues* 19, 368–388.

索　　引

（译名后数字为原书页码,即本书边码）

abstract，抽象的，23，31－32，42，68

Adams，F. 亚当斯，54－56

affect，情感，3，6，30－37，41，42，50，58－60，78－84，95，111

Alexander，J，亚历山大，4，16，20，24，27，28，34，43，46，49，79，94，96，98－99，103，111，135，141

Alicke，M. 艾利克，73

American car case，美国汽车例子，72－73，98

Annas，J. 安纳斯，5

anti-intellectualism，反理智主义，41

Appiah，A. 阿皮亚，5

Arico，A. 阿里科，4

bank cases，银行例子

concrete high standard case，明确的高标准例子，45－46

DeRose bank cases，德娄斯银行例子，36－38

high stakes alternative case，风险大且有出错的可能性的例子，40

high stakes case，风险大的例子，41－42

high stakes no alternative case，风险大且没有出错可能性的例子，40
high stakes typo case，风险大的拼写错误的例子，46－47
high standards case，高标准例子，39
important case，重要的例子，43－44
low stakes case，风险不大的例子，41－42
low stakes typo case，风险不大的拼写错误的例子，46－47
low standards case，低标准的例子，38－39
revised low standard case，修改了的低标准的例子，45－46
simplified high stakes case，简化了的风险大的例子，42－43
simplified low stakes case，简化了的风险不大的例子，42－43
unimportant case，不重要的例子，43－44
Baron，J. 巴隆，97
Batson，C. 巴特森，5
Bealer，G. 贝勒，4，11，18－19，21，24，26，28，29
Beebe，J. 毕比，69
Bengson，J. 本森，28，134
Berker，S. 博克，7
blame，责任，39，54－57，59－60，80，81
Boghossian，P. 博加桑，46，84
BonJour，L. 邦久，28
Brown，J. 布朗，59
Buckner，C. 布克纳，94
Buckwalter，W. 布克瓦尔特，38－40，46－47，54，69，74－76
Burra，A. 伯拉，72
button case，按钮的例子，80－81

Cacioppo, J. 卡西欧普, 48
Camp, E. 坎普, 132
Carey, S. 凯瑞, 8
caring lieutenant case, 关切的中尉的例子, 65 - 66
car thief case, 偷车贼的例子, 58 - 59, 80
causal theory of reference, 指称的因果理论, 14 - 16
causation, 因果性, 69
Chalmers, D. 查默斯, 19
character, 性格, 5
child case, 孩子的例子, 75, 102
Chomsky, N. 乔姆斯基, 60, 101
Christensen, D. 克里斯汀森, 46
Cohen, S. 科恩, 48, 49, 113
Cokely, E. 科克雷, 33, 108
compatibilism, 兼容主义, 16, 29 - 36, 41
concepts, 概念, 2 - 3, 6, 8 - 9, 14, 19, 24 - 26, 33 - 35, 50 - 51, 54, 66 - 69, 74, 77, 84 - 85, 87, 90 - 94, 98 - 99, 126, 138, 142
conceptual analysis, 概念分析, 51
conceptual competence/performance, 概念能力/运用, 24 - 25, 32 - 33, 51, 60 - 61, 69, 93 - 94, 98 - 99
concrete, 明确的, 23, 31 - 32, 42, 45 - 46, 61
consciousness, 意识, 4 - 5
contextualism, 语境主义, 36 - 48, 54, 59, 113
contrastivism, 对比主义, 46
conversational implicature, 会话的暗示, 54, 73
conversational pragmatics, 会话语用论, 54 - 57, 111 - 112

Cullen, S. 库伦, 111 - 112, 134
Cumby, J. 坎比, 69
Cummins, R. 卡明斯, 70, 85
Cushman, F. 库什曼, 67 - 68, 72, 111, 133

Darley, J. 达利, 5
DeRose, K. 德娄斯, 36, 49, 54, 113
descriptivist theory of reference, 指称的摹状词理论, 14 - 16, 19
determined killer case, 坚定的杀手, 76
determinism, 决定论, 16, 29 - 36, 39, 41, 78
Deutsch, M. 杜尔歇, 100, 101
Devine, P. 德维尼, 79
Devitt, M. 德维特, 70, 85, 122
Ditto, P. 蒂托, 78
doctrine of double effect, 双重作用学说, 17 - 18, 23
dog case, 狗的例子, 63 - 64, 86
Doris, J. 多丽斯, 5
drunk driver case, 喝醉了的司机的例子,56, 60
dual process theory, 双进程理论, 7

Elga, A. 埃尔加, 46
Elgin, C. 埃尔金, 70, 85
emotionism, 情绪主义, 6
environmental harm case, 危害环境的例子, 52, 54 - 55, 58, 62, 66 - 68, 70, 75
environmental help case, 改善环境的例子, 52, 55, 62, 66 -

68, 70, 74, 75
Ericsson, K. 埃里克森, 91
Erlenbaugh, J. 厄伦巴赫, 26
expertise defense, 专业知识辩护, 90 - 98
explanation, 解释, 8 - 9
extra dollar case, 多掏的美元的例子, 62 - 63, 84

Fantl, J. 方特, 48, 49, 55
fat man case, 胖子的例子, 78, 111
Feldman, R. 费德曼, 46
Feltz, A. 菲尔兹, 33 - 34, 42 - 43, 57, 103
Fields, J. 菲尔德兹, 2
folk psychology, 大众心理, 4, 69, 73
Foot, P. 福特, 23
Frankfurt, H. 法兰克福, 16 - 17
free cup case, 免费的杯子的例子, 62, 85
free will, 自由意志, 16, 29 - 36
Friedman, O. 弗里德曼, 75, 103

Gendler, T. 金德勒, 14
gene, 基因, 8 - 9
Gettier, E. 葛梯尔, 12 - 13
Gettier cases, 葛梯尔案例, 12 - 13, 15, 19, 103 - 104, 140
Glass, A. 格拉斯, 43
Gödel case, 哥德尔案例, 15 - 16, 23, 73 - 74
Goldman, A. 戈德曼, 14, 20, 50, 68, 76 - 77, 86, 101, 125, 138

Gonnerman, C. 贡纳曼, 94
Gray, H. 盖瑞, 4
Gray, K. 盖瑞, 4
Greene, J. 格林, 7
Grice, P. 格莱斯, 73
Griffiths, P. 格里菲斯, 9
Grundmann, T. 古伦德曼, 122
Guglielmo, S. 古勒莫, 60, 81

habitual tax cheat case, 习惯性逃税的例子, 34
Haidt, J. 海德, 6, 97
Hales, S. 海勒斯, 28, 46, 90, 113, 122
Harman, G. 哈曼, 5
harmful terrorist case, 有害的恐怖分子的例子, 64 - 65, 88
Hawthorne, J. 霍桑, 49, 55, 64
helpful terrorist case, 有益的恐怖分子的例子, 65, 88
Henderson, D. 赫德森, 24
Hintikka, J. 辛提卡, 101
Hitchcock, C. 希区柯克, 69
Holtzman, G. 霍尔兹曼, 76
Horgan, T. 霍根, 24
Horvath, J. 霍瓦特, 122
Huebner, B. 胡伯纳, 4
Hull, J. 胡尔, 40

Ichikawa, J. 市川, 100, 101, 132, 138, 142
incompatibilism, 不兼容主义, 16, 29 - 36

intentional action，不故意行为，52-68，70，72，74，76，78，80，81，84，85，86，88，90，129
interest-relevant invariantism，利益相关的不变主义，49
intuitional diversity，直觉多样性，72-77
cultural diversity，文化多样性，72-74
gender differences，性别差异，74-77
intuitional sensitivity，直觉敏感性，77-81
intuitions，直觉，见哲学直觉(philosophical intuitions)条目
Isen，A. 艾森，5

Jack，A. 杰克，4
Jackman，H. 杰克曼，125
Jackson，F. 杰克森，19
Jarvis，B. 贾维斯，132

Kamm，F. 卡姆，23
Kamtekar，R. 坎特卡，5
Kane，R. 凯恩，38
Kauppinen，A. 库皮尼，24-25，96-97，112，134
Kelly，S. 凯利，88
Kelly，T. 凯利，46
Knobe，J. 诺布，4，11，30-34，42，44-47，52-58，60-61，68-70，72-73，76，78，81，84，87，88，93，125
knowledge，知识，12-13，15，19-20，23，28，29，36-48，50-51，54，57，59，61，64，68-69，77，79-80，82，92-93，97，102-103，105，110
Kornblith，H. 科恩布利思，11，14，25-26，34，70，84-85，

92, 94 - 96, 131
Kripke, S. 克里普克, 15, 73 - 74
Kupperman, J. 库普曼, 5

Lehrer, K. 莱勒, 14, 79
Leslie, A. 莱斯利, 72
Levin, J. 列文, 19
Levin, P. 列文, 5
Lewis, D. 刘易斯, 20, 49, 103, 113, 132
Liao, M. 里奥, 11, 23, 111
Livengood, J. 列文古德, 100, 130
Lombrozo, T. 罗布伦佐, 8 - 9
Ludwig, K. 路德维希, 22, 24, 31, 90, 99, 100, 122, 126, 129
Lynch, M. 林奇, 22, 24, 95, 115

McCann, H. 麦肯, 72
McGrath, M. 麦克格雷斯, 48, 49, 55
Machery, E. 马舍雷, 4 - 5, 54, 60 - 64, 73, 100
Malle, B. 马勒, 56, 60, 72, 76, 78, 81
Mallon, R. 马伦, 64 - 65, 73, 87, 88
Marr, D. 马尔, 43
Marti, G. 马蒂, 100
May, J. 梅, 40, 46
Mele, A. 米尔, 67 - 68, 72, 76
Mendlow, G. 曼德卢, 72, 81
Milgram, S. 米尔格兰, 5

Miller, D. 米勒, 79
Molyneux, B. 莫利努克斯, 26
Monin, B. 莫尼, 79
moral judgment, 道德判断, 6 - 7, 54, 57, 60 - 61, 73, 111
moral psychology, 道德心理学, 5 - 8
moral responsibility, 道德责任, 12, 16 - 17, 28 - 36, 39, 41, 42, 45, 47 - 48, 78, 82
Morris, S. 莫里斯, 29
multiple concepts view, 多种概念观点, 34 - 35
murderous husband case, 蓄意谋杀的丈夫的例子, 32, 42
Murray, D. 默里, 47

Nadelhoffer, T. 纳德尔霍弗, 11, 29, 33, 54, 57 - 61, 72, 73, 78, 80
Nado, J. 纳都, 73
Nagel, J. 内格尔, 14, 61
Nahmias, E. 纳米亚斯, 11, 29 - 30, 32, 39, 41, 47
Nelson, S. 尼尔森, 78
neuroethics, 神经伦理学, 7
Nichols, S. 尼科尔斯, 6, 11, 30 - 34, 41 - 42, 54, 66 - 67, 72, 74 - 75, 78, 93, 102, 104, 106, 111, 125
Nickerson, R. 尼克尔森, 97
Nisbett, R. 尼斯比特, 24, 96, 97, 122
Norenzayan, A. 诺伦泽杨, 3, 97
normative, 规范的, 6 - 7, 17, 19, 35, 51 - 52, 57, 66, 68 - 69, 84, 113
Norton, M. 诺顿, 79

no skill/immoral outcome case, 没有技巧/不道德的结果的例子, 53, 56, 72
no skill/positive outcome case, 没有技巧/正面结果的例子, 53, 72

O'Connor, T. 奥康纳, 38
O'Neill, P. 欧尼尔, 80 - 81, 111
order effects, 顺序效应, 3, 79 - 81, 110

Pereboom, D. 佩雷布姆, 38
performance error, 执行错误, 见概念能力/概念运用(conceptual competence/performance)条目
Petrinovich, L. 彼得诺维奇, 80 - 81, 111
Pettit, P. 佩蒂特, 68, 73
Petty, R. 佩蒂, 110
Phelan, M. 费伦, 43 - 44, 65 - 66, 84, 87, 90
philosophical analysis, 哲学分析, 2, 16, 28 - 49
philosophical intuitions, 哲学直觉, 11 - 27,89 - 108
  doxastic conception, 信念观念, 20 - 21
  etiological conception, 病原学观念, 24
  methodological conception, 方法论观念, 25
  phenomenological conception, 现象学观念, 21 - 23
  semantic conception, 语义学观念, 23
philosophical methodology, 哲学方法论, 3, 19 - 20, 70 - 89, 101, 109, 111, 110
philosophy of mind, 心灵哲学, 3, 50 - 69, 84
Pinillos, N. A. 皮内罗斯, 44, 46 - 47, 64

Pink, T. 平克, 117 n. 1
Pizarro, D. 皮萨罗, 78
Plant, E. 普兰特, 79
Plantinga, A. 普兰廷加, 28
police officer case, 警察的例子, 58 - 59, 80
pragmatics, 语用学, 见会话语用学(conversational pragmatics)
praise, 赞扬, 39, 74
principle of alternative possibilities, 另一种可能性原则, 16 - 17
Prinz, J. 普林茨, 4, 6
Pronin, E. 普罗宁, 97
push case, 推的例子, 80 - 81, 111
Pust, J. 普斯特, 22, 26, 68, 86
Pylyshyn, Z. 派利夏恩, 43

reflection, 反思, 7 - 8, 24 - 25, 96 - 97, 107
reflective equilibrium, 反思平衡, 122
relativism, 相对主义, 35, 46, 77, 113
restrictionist challenge, 限定主义质疑, 70 - 89
Robbins, P. 罗宾斯, 4
Rorty, R. 罗蒂, 113
Roskies, A. 罗斯基, 41
Ross, L. 罗斯, 2
Roxborough, C. 洛克希布劳, 69
Rust, J. 鲁斯特, 8

salience, 显著性, 29, 36 - 48, 50, 52, 59, 61

Sarkissian, H. 萨尔吉希安, 65-66, 84, 87, 90
Schaffer, J. 谢弗, 37, 44-47
Schultz, E. 舒尔兹, 133
Schuman, H. 舒曼, 2
Schwarz, N. 史怀泽, 111
Schwitzgebel, E. 史怀泽格贝尔, 8, 111, 133
sentiments, 情感, 6
serial rapist case, 强奸惯犯的例子, 34
Shanteau, J. 桑蒂, 91
side-effect effect, 副作用效应, 52-69, 76, 84
Simpson, S. 辛普森, 64
Singer, P. 辛格, 7, 91, 122
Sinnott-Armstrong, W. 辛诺特-阿姆斯特朗, 40, 110
situationism, 情境主义, 5
skepticism, 怀疑主义, 21, 85-87, 102-103, 105, 107, 111, 145
Sommers, T. 索默斯, 11, 47
Sosa, E. 索萨, 14, 23-24, 26, 28-29, 83, 96-97, 110
Sreenivasan, G. 斯里尼瓦萨, 5
stakes, 风险, 29, 36-48, 50, 54, 57, 59, 63, 64
Stalnaker, R. 斯塔纳克, 19
Stanley, J. 斯坦利, 41, 49, 64
Starmans, C. 斯塔曼斯, 75, 103
Steadman, A. 斯特德曼, 54-56
Stich, S. 斯蒂克, 72, 74-76, 122
Stotz, K. 斯托兹, 9
stranger case, 陌生人的例子, 74-75, 102

Strawson, G. 斯特劳森, 38
subject-sensitive invariantism, 主体敏感的不变主义, 36－48, 59
supercomputer case, 超级计算机的例子, 29－30, 39
Swain, S. 斯温, 79
switch case, 手表的例子, 80－81, 111
Sytsma, J. 斯特斯马, 4, 100

Talbot, B. 塔尔伯特, 28
Tannenbaum, D. 坦南鲍姆, 78
thickness defense, 厚辩护, 99－100
Thomson, J. 汤姆森, 17－18, 23, 78
trade-off hypothesis, 交易假说, 62, 64, 84, 90
transplant case, 器官移植的例子, 18
trolley problem, 电车难题, 7, 17－18, 23, 63－64, 74－75, 78－79, 80－81, 102, 111
Truetemp case, 特鲁特普案例, 13－14, 79－81, 110
Tsu Ch'ung Chih case, 祖冲之案例, 73－74
Turner, J. 特纳, 29

Uhlmann, E. 乌尔曼, 78, 107, 108
Ulatowski, J. 乌拉托夫斯基, 54, 66－67

van Inwagen, P. 范·因瓦根, 21

Warfield, T. 沃菲尔德, 46
Watson, P. 沃森, 97

watch case，手表的例子，75
Weatherson, B. 威瑟森，15，64
Wegner, D. 温格纳，4
Weinberg, J. 温伯格，11，16，20，28，34，72，77，79，83，91，94，96，98-99，110-111，122，131，135，141
White, R. 怀特，46
Williamson, T. 威廉姆森，14，21，26，27，29，97，102-105，122，130，140，142，143
Wilson, T. 威尔森，24，96，131
worker case，工人的例子，63-64，86
Wright, J. 赖特，110

Young, L. 杨，72

Zamzow, J. 赞姆佐，74-75，102，104，106，111
Zarpentine, C. 扎彭提尼，42-43，57
Zimmerman, A. 齐默尔曼，40

译 后 记

一篇译后记的好处在于，我可以在这里说一些不可能包含在任何“学术文章”之内的东西，然而它们的重要性，至少对我而言，却高于学术文章。

国内每年都会出版很多译著，哲学类的译著当然也不少，因此这本《实验哲学导论》只是“很多本”中的一本，它的意义（对他人的影响）不取决于书本身，也不取决于我，因此对我来说它有没有意义都是无所谓的。它很可能没有什么意义，它浪费了很多纸张，它耗损了编辑王巧贞同志的脑力和视力，它折磨着校对工作者，它让那些印刷工人汗流浃背地工作，他们被某种我说不上来的化学物质熏坏了眼睛，它让那么多的人为它受苦。而它却很有可能和很多其他出版物一样，很快就被人遗忘了。甚至不能说被人遗忘，因为它可能从未被人记住（这同样适用于当今速朽的“文化产业”）。

这让人绝望。裴多菲有这么一句话：“绝望之为虚妄，正如希望相同。”从前我很不能理解它，现在也不能说理解了，不过终究有了一个解释，那就是人的缺省状态是不愿意绝望的，是希望有希望的。但人之所以有希望，一定是什么东西给人以希望了，而这“什么东西”很可能是虚假的，或很快就会被收回，因此希望也是虚妄，绝望继续。然而，个人又无法改变绝望的处境，因此绝望也是虚妄。既然都是虚妄，那怎么办呢？答案也许是：继续

做。并且首先不是为了意义而做，那是别人的事，首先要为自己而做。

2012年的夏天，厦门大学知识论与认知科学研究中心的曹剑波、郑伟平和我刚刚看完另一本书《实验哲学》的校样。那是一本论文集，现已正式出版了。我以为关于实验哲学的翻译工作可以暂告一个段落了，但是很快就收到了曹剑波教授的邮件，他"建议"我另外再翻译一本书（作为我的顶级老大哥，他的"建议"庶几等于命令）。这本名为《实验哲学导论》的书出版于2012年，算是新鲜出炉了。当时我没看过这本书，也想不好要不要译它，就抱着"聊聊"的心态，给作者约书亚·亚历山大写了封邮件，后者呈现出了一个年轻学者特有的温和、热情，并且自动把我的"聊聊"变成了"干吧"。一来没人能抵挡得了热情和温和，二则既然已经进入了人与人的交往阶段，再说不干，恐怕就有点说不过去了。

就这样，恰恰因为这两个人，也仅仅因为这两个人，我便着手翻译这本书了。我躲在厦门大学人文学院的一间小小的没有窗户的办公室里——没有窗户其实是件好事，因为如果我的办公室有窗户，那么它必定正对着楼下的上弦场（一个操场），那么我一趴在窗户上就能看见那些永远百无聊赖，永远着装怪异，永远大惊小怪，永远在摆弄手机、相机的游客们，我不愿意见到他们——搞我的翻译。顺便说一下，厦门大学是一座坐落于海边的常年鲜花盛放的大学，这"鲜花盛放"本身就足以使其脱离时下流行的大学评价机制而成为一所优秀的大学了，因为所谓的机制总是会变，而且确实经常改变，而南国的植物却如同真理一般古老而坚定，这是我喜欢她的最重要的原因。

翻译当然是一件吃力不讨好的活儿，但也有它特有的乐趣，那就是遇到了不懂的地方（应该说本书并不难懂，但由于书中的

术语和观念都很新,导致我确实有很多不懂的地方,有几处现在仍然存留着一点疑问),或查找其他文献,或自己琢磨,或询问作者,最后将其搞懂,这是很愉快的,像是回到了解几何难题的初中时代。我前面说了,做事情不是为了意义,而是为了自己,因此我在搞翻译的时候首先要求自己先看懂原文,明白了作者的意思,然后再译。“先看懂,再翻泽”这个要求貌似不是要求,貌似是翻译的基本条件,其实是很难做到的。由于各种原因,我(们)看不懂的情况是有的,而我(们)总是倾向于“不懂就译”,借助字典,外带一点语感,把一个词或一句话译成中文,那就算完了。虽然看懂后再翻译和不搞懂就翻译出来的句子可能是一样的,但是你要遇到译者,让他解释清楚某些你不懂的地方,那是不可能的。幸运的是,同时也是不幸的是,我们一般不可能遇到译者,因而这事也就稀里糊涂地过去了。

在本书中就有好几处这样的地方。

比如文中“引言”一章内有“现象意识(经验)”和“非现象意识(主体)”的说法,现象意识的原文为“phenomenal consciousness”,经验为“experience”,非现象意识的原文为“non-phenomenal consciousness”,主体为“agency”。从英文翻译成中文,是挺简单的,似乎也只能这样翻译了。但是现象意识和非现象意识这两个概念,我是不懂的,查了若干篇英文文献之后,也只是大体上搞懂了。现象意识就比如“看到红色”、“感到痛”、“听到一个声音”,这些“意识”是经验性的(感觉、视觉、听觉在起作用),而“具有某个信念或愿望”之类的,则是非现象意识,很难说有什么经验与其关联。而原文中之所以有“agency”(主体)一词,乃是因为哲学家一般认为现象意识是接受性的,而非现象意识是(与接受性相对立的)主动性的,是主体的行为。这就基本清楚了,翻译下去,也就不心虚了。

再比如,原文第二章中有“compute the function”的说法,我将其翻译成了“计算某个函数”。似乎也只能翻译成这样了。对于认知科学,我也是不懂的,查找了一些文献,也详细询问了作者,得出结论:这个 function 不能翻译成“功能”,而只能翻译成“函数”。这是认知科学的一种说法,根据心灵的运算理论,人类的心灵或大脑是一个信息处理系统,大脑中的各个“机制”是一个个能够独立运算的“模块”。我们容易知道输入和输出(输入某个模块,从这个模块中输出)是什么,但是对大脑作为某个模块所计算的函数本身是什么,则是不清楚的——这应该是作者想要表达的意思。

另外,文中一些我自己不太懂的地方,或者我非法地猜测读者可能会有疑问的地方,或者粗暴地臆想读者可能会产生误解的地方,或者有些英文词已经有了约定俗成的译法而我又不太满意的地方,我均为其添加了“译者注”——尽管译者注也会让人产生进一步的疑问或误解,这就不去管它了。因此文中有很多译者注,可能会显得甚为啰嗦,这是要请那些学识广博的读者原谅的,因为对已经懂这些问题的读者而言,译者注当然是画蛇添足。不过要没有这些“添足”,我就不舒服。举个例子,原文中经常出现这样一个词——Side Effect,这个词常被译成“边际效应”或“边界效应”,而我把它译成“副作用”,因为(至少在我看来)“边际效应”不清晰、不明白,至少不如“副作用”那样清晰、明白。Side Effect 说的是:我想做一件事,我想要这事带来某种影响,这种影响是主要的,是合乎我的目的的,然而这事还会带来一些另外的影响,这种影响是次要的,我也能预期到这个影响,就像药品的“副作用”,治病这回事是主要的,它一定得发生(治不治得好那是另一回事),副作用则是次要的。

这样的地方挺多,这里就不赘述了。我想,只要是人,以为

自己懂了而最后却不懂的情况必定是存在的，所以文中一定有很多的差错，请读者朋友不吝赐教。

关于翻译本身，也简单地谈一点心得。有一个词，叫做“译文体”，说的是翻译外语著作的汉语总有一点不同于作为母语写出来的汉语。如果说翻译界已经有了一种明显的“译文体”，好像不能这样说，但是如果说完全不存在“译文体”，似乎又不对。一方面，原著者的文字风格是不同的，另一方面，译者的汉语表达习惯也是不同的，导致即使有所谓的“译文体”，也是五花八门、形形色色（也许这就是翻译吸引人的地方吧），所以，要明确说清楚“译文体”到底是什么，好像也很难。

但是，在翻译中总存在着一个主要矛盾，那就是“原味”和“化归”的矛盾：要保持原味，就得保持原文的句法、结构，就有可能是硬译甚至死译，有可能难读、难懂，不够“达”了（所谓的“信、达、雅”），要“化归”，要绝对的流畅和通达，那就有可能译出的东西像是自己写的，这样的译文看起来固然爽快，但是经验证明这种做法是不对的，而且译文肯定是不“信”，甚至有遗漏的。翻译家们当然能够很好地处理这一矛盾，但我是肯定做不到的。不过，因为不愿过分“化归”，所以我译出的东西死硬而呆板，可能要让读者如梁实秋那样伸出手指、硬着头皮来读了。——没办法，这也是要请读者原谅的。读者若有机会和精力看英文原版，那当然最好，那本书就可以凤凰涅槃了。若在涅槃之前，还能让读者有一点所得，那我也就欢喜异常了。

我要由衷地感谢我在厦门大学哲学流动站做博士后期间的指导老师陈嘉明先生，还有厦门大学哲学系的曹志平主任，没有他们的教诲和帮助，我是不可能从事哲学工作的；感谢我的顶级老大哥——厦门大学哲学系的曹剑波和郑伟平，他们为我提供了不可替代的帮助；还要感谢（这“感谢”还伴随着一些特有的经

验，因此它稍微不同于其他的“感谢”）我的妻子张娜，近一年多来，她义无反顾地承担着繁重的家务劳动，还悉心照料着一个尚处“小动物”阶段，却每每令人疲惫不堪的婴儿。

最后，我还要感谢本书作者约书亚·亚历山大，他极其耐心地解答了我的各种疑问；感谢我的编辑王巧贞细致而认真的工作。我和约书亚、王巧贞之间极为频繁的邮件通信，已经构成了一张密不透风的网。本书的出版，就是这张网的结果。

二十世纪西方哲学译丛

卡西尔:《人论》《人文科学的逻辑》
弗洛伊德:《弗洛伊德后期著作选》
胡塞尔:《逻辑研究》(修订本)《生活世界现象学》《现象学的方法》
　　　　《欧洲科学危机和超验现象学》
柏格森:《创造进化论》
波普尔:《客观知识》《猜想与反驳》
卡尔纳普:《世界的逻辑构造》
拉卡托斯:《科学研究纲领方法论》
蒯因:《从逻辑的观点看》
艾耶尔:《语言、真理与逻辑》《二十世纪哲学》
加达默尔:《真理与方法》《哲学解释学》
马尔库塞:《爱欲与文明》《单向度的人》
罗素:《我们关于外间世界的知识》
赖尔:《心的概念》
克里普克:《命名与必然性》
萨特:《存在主义是一种人道主义》
雅斯贝斯:《时代的精神状况》《生存哲学》
法伊尔阿本德:《反对方法》《自由社会中的科学》
本格:《科学的唯物主义》
劳丹:《进步及其问题》
夏佩尔:《理由与求知》
普特南:《理性、真理与历史》
塞尔:《心灵、语言和社会》《心、脑与科学》
米德:《心灵、自我与社会》
沃特金斯:《科学与怀疑论》
莱维-斯特劳斯:《结构人类学》
奥伊肯:《生活的意义与价值》
德里达:《论文字学》
T・内格尔:《人的问题》
E・内格尔:《科学的结构》
范・弗拉森:《科学的形象》
欧克肖特:《政治中的理性主义》
海德格尔:《林中路》(修订本)《现象学之基本问题》《物的追问》《康德与形而上学疑难》
利科:《历史与真理》《活的隐喻》
罗蒂:《后哲学文化》
帕菲特:《理与人》
阿佩尔:《哲学的改造》
达米特:《分析哲学的起源》
巴雷特:《非理性的人》
丹图:《叙述与认识》
韦尔默:《后形而上学现代性》
罗斯:《正当与善》
马里翁:《还原与给予》
诺奇克:《合理性的本质》
亚历山大:《实验哲学导论》

封面
书名
版权
前言
目录
中文版序言
致谢
引言
第一章　哲学直觉
第二章　实验哲学和哲学分析
第三章　实验哲学和心灵哲学
第四章　实验哲学和哲学方法论
第五章　为实验哲学辩护
后记
参考文献
索引
译后记